AD-A075 549   STANFORD UNIV CA DEPT OF OPERATIONS RESEARCH   F/G 12/1

SIMULATING GENERALIZED SEMI-MARKOV PROCESSES.(U)

SEP 79   L D FOSSETT   N00014-76-C-0578

UNCLASSIFIED   TR-50   NL

1 OF 2

AD-
A075549

LEVEL

12

6

SIMULATING GENERALIZED SEMI-MARKOV PROCESSES*

by

10 Lawrence D. Fossett

9 TECHNICAL REPORT NO. 50

11 September 1979

12 103

15

Prepared under Contract N00014-76-C-0578 (NR 042-343)
NSF-MCS75-23607
for the
Office of Naval Research

14 TR-50

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

*This research was also partially supported under National Science Foundation Grants MCS75-23607 and MCS-7909139.

50B

402766 79 10 17 064

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

IN RECENT YEARS the use of queueing networks as models for complex stochastic systems has expanded dramatically. In computer modeling, for example, queueing networks have been used as models of data management systems, systems with many peripheral processors and demand paging systems, among many others. As the models become more detailed, analytical results become very difficult to obtain. For this reason, simulation has become an increasingly attractive avenue for the study of these models.

Typically, the goal of the simulation is the estimation of some characteristic of the system under study. However, determining how close the simulation's estimates are to the correct value is a major problem. A common approach to this difficulty is the use of confidence intervals. This method is most fruitful when the variance of the estimate is easily obtained; but, unfortunately, this is often a difficult problem.

When the characteristic under study is a function of the long run 'equilibrium' behavior of the system, a special class of processes called regenerative processes can sometimes be used to greatly simplify the problem of variance estimation. These processes periodically 'begin from scratch' in the sense that at certain epochs in time, the system behaves as it did originally. This allows

each sample path to be broken into independent and identically distributed pieces, which makes the variance calculation straightforward. It is therefore important to find ways to model queueing networks as regenerative processes. This is our first goal.

One approach to modeling queueing networks and other complex stochastic systems which has received some attention in the literature is the generalized semi-Markov process (GSMP). This idea is an example of the supplementary variables approach to non-Markovian systems described in Cox and Miller (1966). This approach 'supplements' the natural description of the system by variables which contain information about the past history of the system. In this way, a model of a non-Markovian system can be made Markovian. For GSMPs the supplementary variables are clocks which record the amount of time until the occurrence of various events which could influence the system. In a queueing network, for example, each server and each arrival stream would be associated with a clock. By including these clock readings as part of the description of the system, only the present state of the system is required to predict future behaviour. This means that the new model is Markovian and therefore amenable to analysis via the use of Markov chain theory.

To use these processes for simulation purposes a central limit theorem is required. Obtaining this result is the second goal of this paper. Our approach to this problem is to find closely related regenerative processes on which to base the central limit theorem for the process under study. New results in the theory of Markov chains on a general state space make it clear how these regenerative processes can be constructed.

A natural starting point for obtaining our objectives is a review of the theory of Markov chains. Section 2.1 contains a review of the applicable limit theorems. The hypotheses for these results have developed as various

recurrence conditions and these are also reviewed. The structure and notation of generalized semi-Markov processes are introduced in Section 2.2. The final section of Chapter II is a brief summary of the regenerative approach to simulation.

The main results of this paper are presented in two chapters. The first section of Chapter III identifies a class of GSMPs which are regenerative. If a queueing network occupies a state where only one event can change the state of the system, (when all jobs in a closed network wait for a single server, for example) the process behaves the same way each time it leaves this state. This phenomenon can be observed in any GSMP that has a single event associated with any state. This is the critical feature of regenerative processes. Section 3.2 presents two examples where regenerative methods are used in the simulation of queueing networks.

For those GSMPs without a 'single' state, Chapter IV describes how results of Markov chain theory can be applied to obtain central limit theorems for functions of the 'equilibrium' behavior of the GSMP. The results in Section 4.2 determine which GSMPs satisfy the recurrence conditions introduced in Section 2.1. As long as the clocks governing the events of the GSMP have bounded positive densities on the positive half line, at least the weakest forms of recurrence are satisfied. Section 4.3 presents an ergodic theorem for these GSMPs. A large class of these recurrent GSMPs can be associated with a regenerative process by a further use of the supplementary variables technique. By supplementing the GSMP with a 'memory' of a finite number of steps, an iterate of the transition function of the GSMP can be split into a state-invariant measure and a state-dependent remainder. By replacing this iterate of the transition function by a judicious choice between the invariant measure and the remainder a regenerative process can be formed which has marginal distributions identical to the GSMP's. Unfortunately, this

regenerative process is not suitable for simulation applications because the the remainder cannot be determined explicitly in most cases. To circumvent this problem, we propose a new supplementary variables technique which only requires a sequence of Bernoulli random variables. The development of this procedure and the central limit theorem for the new supplemented process is discussed in Section 4.4. Estimation of the central limit theorem's variance constant is the topic of Section 4.5. An example testing extensions of the theory is presented in Section 4.6.

Chapter V examines the strengths and weaknesses of this new technique.

Before beginning, it is convenient to define some notation which will be used throughout our discussion. The indicator function $1_B(x)$ is 1 or 0 according as $x \in B$ or $x \notin B$. Then if X is a random variable, $1_B(X)$ is also, and equals 1 if the event $\{ X \in B \}$ occurs and 0 if it does not. We will use R^n to denote the cross product of n copies of the real numbers. If n is omitted, it will be assumed to be 1.

All results with original proofs will be labelled theorems, while results cited from other sources will be called lemmas. The symbol $\square$ will denote the conclusion of a proof.

CHAPTER II

PRELIMINARIES

2.1. MARKOV CHAINS ON A GENERAL STATE SPACE

IN THIS SECTION notation and some limiting results are introduced for Markov chains on a general state space. These results are useful is the development of central limit theorems for GSMPs. Most of these theorems demand that the chain satisfy some kind of recurrence property and a review of these conditions is an important part of the discussion. For a more complete discussion of these notions, see Neveu (1984), Orey (1971) and Revuz (1975). Throughout this section (E,$\mathcal{E}$) will be a measurable space.

(1) DEFINITION. A function $P:(E,\mathcal{E}) \to [0,1]$ is said to be a probability transition function (or kernel) if:

(a) $P(x,\cdot)$ is a probability measure on (E,$\mathcal{E}$) for all $x \in E$,

(b) $P(\cdot,B)$ is a measurable function with respect to $\mathcal{E}$, for all $B \in \mathcal{E}$.

The n-step transition probabilities are defined by setting $P^1(x,B) = P(x,B)$ and $P^{n+1}(x,B) = \int_E P^n(x,dy)P(y,B)$.

In the usual fashion, let $E^\infty = E \times E \times E \times \cdots$ and $\mathcal{E}^\infty = \mathcal{E} \times \mathcal{E} \times \cdots$. For $\omega \in E$, let $X_i(\omega)$ be the i^{th} coordinate of ω. For any probability measure ν on $(E,\mathcal{E})$, there is a probability measure P on $(E^\infty,\mathcal{E}^\infty)$ satisfying for any

$n > 0$ and $B_1, B_2, \ldots, B_n \in \mathcal{E}$ the following relation:

$$P_\nu(X_0 \in B_0, X_1 \in B_1, \ldots, X_n \in B_n)$$
$$= \int_{B_0} \nu(dx_0) \int_{B_1} P(x_0, dx_1) \cdots \int_{B_n} P(x_{n-1}\, dx_n).$$

The properties of the measure P_ν ensure that conditional probabilities can be constructed which satisfy

$$P_\nu(X_n \in B \mid X_0 = x_0, X_1 = x_1, \ldots, X_k = x_k) = P^{n-k}(x_k, B), \qquad (2)$$

for $n > 0, x \in E, k = 0, 1, \ldots, n.$

Equation 2 is a version of the Markov property, and when it is satisfied $X = \{X_n, n \geq 0\}$ is said to be a Markov chain on state space E with initial distribution ν and stationary transition probability function P.

The notion of recurrence captures the idea of infinitely many visits by the Markov chain to a state or a collection of states. In a countable state space Markov chain, a state is called recurrent if, starting in state i, the chain returns to state i with probability 1. If, in addition, all the states communicate, the chain is said to be recurrent. If, however, the state space of the Markov chain is not countable, it may not happen that a particular state is visited infinitely often; a more general concept is therefore needed. One natural way to generalize the idea is to require that 'significant' sets are visited 'frequently enough.' To study these conditions a measure closely related to P must first be introduced.

(3) DEFINITION. A σ-finite measure π on (E,$\mathcal{E}$) is invariant for the transition probability function P if, for all $x \in E$ and $A \in \mathcal{E}$

$$\pi(A) = \int_E P(x, A)\, \pi(dx).$$

The first general notion of recurrence is a local condition due to Harris (1956) which has recently been reformulated by Athreya and Ney (1978) and Nummelin (1978).

(4) DEFINITION. (a) The chain X is said to be φ-recurrent (or Harris recurrent) if there exists a positive σ-finite, invariant measure φ on E such that $\varphi(A)>0$ implies

$$P\left(\sum_{m=1}^{\infty} 1_A(X_m) = +\infty\right) = 1.$$

(b) A Markov chain X on (E,$\mathcal{S}$) is (A,B,φ,λ,n)-recurrent if there exists sets A and $B \in \mathcal{S}$, a probability measure φ on B, a finite number $\lambda > 0$, and a finite integer n such that

(i) $P_x(X_k \in A \text{ for some } k \geq 1) = 1$ for all $x \in E$; and

(ii) $P_x(X_n \in C) \geq \lambda\varphi(C)$ for all $x \in A$ and $C \subseteq B$.

If either (a) or (b) is satisfied X will be said to be recurrent. If in part (b) A=B, then the chain will be said to be (A,φ,λ,n)-recurrent. (A,φ,λ,n)-recurrence is in fact no less general than (A,B,φ,λ,n)-recurrence. A proof of this fact and of the equivalence of parts (a) and (b) can be found in Athreya and Ney (1977, 1978). Several limit results for these chains are of interest.

(5) LEMMA. For every recurrent Markov chain on (E,$\mathcal{S}$) with transition function P there exists a nontrivial σ-finite invariant measure π. If a measure π' is σ-finite and invariant with respect to P, then π' is a multiple of π. Furthermore, the Markov chain is π-recurrent.

Proof. See Orey (1971). ▯

(6) LEMMA. Let X be recurrent with invariant measure π. If f and g are measurable functions satisfying $\int_E |f|\, d\pi < +\infty$, $\int_E |g|\, d\pi < +\infty$, and $\int_E g\, d\pi > 0$, then

$$\lim_{n\to\infty} \frac{E_x[\sum_{k=0}^{n} f(X_k)]}{E_x[\sum_{k=0}^{n} g(X_k)]} = \frac{\int_E f\, d\pi}{\int_E g\, d\pi}$$

for $x \in E$.

Proof. See Revuz (1975). ∎

(7) LEMMA. Let X be recurrent with invariant measure π. If f and g are measurable functions satisfying $\int_E |f|\, d\pi < +\infty$, $\int_E |g|\, d\pi < +\infty$ and $\int_E g\, d\pi > 0$, then

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n} f(X_k)}{\sum_{k=0}^{n} g(X_k)} = \frac{\int_E f\, d\pi}{\int_E g\, d\pi} \quad P_\nu - a.e.$$

for any probability measure ν.

Proof. See Revuz (1975). When $g = 1_E$, this is the strong law of large numbers. ∎

If X is $(A, \varphi, \lambda, 1)$-recurrent, then condition (ii) of Definition 4(b) allows the decomposition of the transition function P by

$$P_x(x, B) = p\varphi(A \cap B) + (1-p)Q(x, B) \tag{8}$$

for $0 < p \leq \lambda$. The following critical lemma makes clever use of this fact. The idea presented in the proof is integral to the future treatment of the problem, so a sketch of it is included.

(9) LEMMA. If X is an $(A, \varphi, \lambda, 1)$-recurrent Markov chain then there exists a random time N such that $P_x(N < +\infty) = 1$ and

$$P_x(X_n \in B, N = n) = \varphi(B \cap A) P_x(N = n)$$

for all $x \in E, B \in \mathcal{S}$.

Proof. Let $E' = E \times (0,1)$, $\mathcal{S}'$ be the σ-algebra of subsets of E' given by $\mathcal{S} \times \{0,1\}$, and P' be a transition function on $E' \times \mathcal{S}'$ defined by

$$P'((x,\delta),(B,0)) = \begin{cases} (1-p)P(x,B), & \text{if } x \notin A, \\ (1-p)Q(x,B), & \text{if } x \in A; \end{cases}$$
$$P'((x,\delta),(B,1)) = \begin{cases} pP(x,B), & \text{if } x \notin A, \\ p\varphi(A \cap B), & \text{if } x \in A; \end{cases} \tag{10}$$

where $\delta = 0$ or 1 and Q is defined in (8). Let $X' = \{(X_n, \delta_n), n = 1, 2, \dots\}$ be a Markov chain on $(E', \mathcal{S}')$ with transition function P'. Then it can be checked that

(i) $\{X_n, n \geq 0\}$ is a Markov chain on $(E, \mathcal{S})$ with transition function P;

and

(ii) $\{\delta_n, n \geq 0\}$ is a sequence of independent and identically distributed Bernoulli random variables with parameter p.

Define $N = \inf(n \geq 1 : X_{n-1} \in A, \delta_n = 1)$. It remains only to observe that the probability the Bernoulli variables are zero each time the process enters A is zero. ∎

It should be noted that the transition function in (10) is not the same as that reported in Athreya and Ney (1977, 1978). The author was notified of the changes in the transition function by private communication.

An immediate corollary is that a sequence of these times can be found. Unfortunately, only very special processes are $(A, \varphi, \lambda, 1)$-recurrent, but any (A, φ, λ, k)-recurrent process can be closely linked to one of them. There is one additional, but familiar, requirement. For each A, define the set

$$\kappa_A = \{k \geq 1: \text{there exists } \lambda_k > 0, \text{ and a probability distribution } \varphi_k \text{ on A with the property that } P^k(x,B) > \lambda_k \varphi_k(B) \text{ for all } x \in A \text{ and } B \subseteq A.\}$$

If the greatest common divisor of κ_A is 1, the chain X is said to be aperiodic.

(11) LEMMA. Let X be aperiodic and (A, φ, λ, k)-recurrent. Then the skeleton process $Y = \{ Y_n = X_{kn}; n = 0, 1, \ldots \}$ is $(A, \varphi, \lambda, 1)$-recurrent.

Proof. See Athreya and Ney (1978a). ∎

The second type of recurrence is a global condition.

(12) DEFINITION. DOEBLIN'S CONDITION–Doeblin (1940). There exists an positive integer n, two real numbers $\gamma < 1$ and $\eta > 0$, and a probability measure τ on $(E, \mathcal{E})$ such that, for $A \in \mathcal{E}$,

$$\tau(A) > \gamma \text{ implies } \mathrm{P}^n(s, A) > \eta \text{ for all } x \in E.$$

Another formulation of this idea is

(13) DEFINITION. DOOB'S CONDITION–Doob (1953). There is a finite valued measure φ on set $A \in \mathcal{E}$, with $\varphi(E) > 0$, an integer $k > 1$ and a positive ϵ such that, for all $x \in E$,

$$\mathrm{P}^k(x, A) < 1 - \epsilon \text{ whenever } \varphi(A) < \epsilon.$$

Occasionally, the parameters satisfying the conditions will be made explicit by stating a 4-tuple (τ, n, γ, η) for Doeblin's Condition and a triple (φ, k, ϵ) for Doob's Condition. Note that φ can be replaced by a probability measure $\varphi^*(A) = \varphi(A)/\varphi(E)$. If (φ, k, ϵ) satisfies Doob's Condition, then (φ^*, k, δ) does as well, with $\delta = \min[\epsilon/\varphi(E), \epsilon]$. In this report, the measure φ will always be a probability measure.

The literature is somewhat ambiguous regarding the relationship between the two global conditions (see Doob (1953) and Athreya and Ney (1978)). The following theorem clarifies the situation.

(14) THEOREM. Let X be a Markov chain on $(E, \mathcal{E})$. Doob's Condition is satisfied if and only if Doeblin's Condition is satisfied.

Proof. Suppose Doob's condition is satisfied by (φ, k, ϵ). For $B \in \mathcal{E}$ satisfying $\varphi(B) > 1-\epsilon$, or $\varphi(B^c) < \epsilon$ it must be that $P^k(x, B^c) < 1-\epsilon$ or $P^k(x, B) > \epsilon$. Thus Doeblin's condition is satisfied with 4-tuple $(\varphi, k, 1-\epsilon, \epsilon)$.

Conversely, suppose Doeblin's condition is satisfied by (τ, n, γ, η). If $B \in \mathcal{E}$, with $\tau(B) < 1-\gamma$, the condition implies $P^n(x, B) < 1-\eta$. This means Doob's condition is satisfied by triple (τ, n, δ) with $\delta = \min[1-\gamma, \eta]$. ∎

It is interesting to note that X is (E, φ, λ, n)-recurrent if and only if X satisfies Doeblin's condition (See Neveu, 1964). For this reason it is not surprising that the limit results for these chains are stronger than those for Harris recurrent chains. To explore these results more terminology is required.

(15) DEFINITION. A set B will be called a consequent set if, for some $x \in E$, $P^n(x, B) = 1$ for all n, and in this case B will be called a consequent of x. A set which is a consequent of every one of its points will be called an invariant set.

Doob has demonstrated that when his condition is satisfied, there is a decomposition of E into disjoint invariant sets $E_1, E_2, \ldots$ and a transient set $F = E - \bigcup_{i=1}^{\infty} E_i$. Corresponding to each E_i is a probability measure π_i with the properties that

$$\pi(E_i) = 1 \quad \text{and} \quad \lim_{n \to \infty} \frac{\sum_{m=1}^{n} P^m(x, C)}{n} = \pi_i(C) \quad \text{for} \quad x \in E_i, C \subseteq E_i.$$

The E_i are called ergodic sets. Furthermore each E_i can be decomposed into d_i disjoint sets $D_{i,1}, D_{i,2}, \ldots, D_{i,d_i}$ such that

$$P(x, D_{i,j}) = 1 \text{ for } x \in D_{i,j-1} \quad j = 1, 2, \ldots, d_i.$$

The sets $D_{i,j}$ are called cyclically moving subsets of the ergodic set E_i.

(16) **DEFINITION.** If the Markov chain X satisfies Doob's Condition and has a single ergodic set with no cyclically moving subsets ($d_1 = 1$) then condition D_0 is satisfied.

Geometric convergence to the invariant measure is the result of interest for these Markov chains.

(17) **LEMMA.** Let X be a Markov chain satisfying Condition D_0, with an invariant measure π. For some $n_0 < +\infty$, some $\rho < 1$, $n > n_0$,

$$|P^n(x, A) - \pi(A)| < \rho^n$$

where $A \in \mathcal{S}, x \in E$.

Proof. See Doob (1953). ∎

Occasionally it is convenient to redefine Condition D_0 solely in terms of the transition function P. One way to do this is via the coefficient of ergodicity.

(18) **DEFINITION.** Let $(E, \mathcal{S})$ be a measurable space and P be a transition probability function defined on $(E, \mathcal{S})$. The real number

$$\alpha(P) = 1 - \sup_{\substack{x, y \in E \\ B \in \mathcal{S}}} |P(x, B) - P(y, B)|$$

is called the coefficient of ergodicity.

Dobrušin(1956) notes that a necessary and sufficient condition for $\alpha(P)$ to be zero is the existence, for any positive ϵ, of $x, y \in X$ with measures $P(x, \cdot)$ and $P(y, \cdot)$ concentrated on sets B_1 and B_2 with the property that $P(x, B_1 \cap B_2) < \epsilon$ and $P(y, B_1 \cap B_2) < \epsilon$.

For a Markov chain a coefficient of ergodicity can be determined for each iterate of the transition probability function. When one of those coefficients is positive, we shall say that the Ergodic Condition is satisfied.

(19) THEOREM. Let X be a Markov chain on (E,$\mathcal{E}$) with transition function P. Then Condition D_0 is satisfied by X if and only if the Ergodic Condition is satisfied.

Proof. Suppose $\alpha(P^k) > a > 0$. Choose $x \in E$ and any set $A \in \mathcal{E}$ with $P^k(x, A) < a/2$. For any $y \in X$,

$$|P^k(y, A) - P^k(x, A)| < \sup_{\substack{x,y \in X \\ B \in \mathcal{E}}} |P^k(y, A) - P^k(x, A)| = 1 - \alpha(P^k).$$

Therefore,

$$P^k(y, A) < P^k(x, A) + (1 - a) < 1 - \frac{a}{2}.$$

This means that Doob's Condition is satisfied by triple $(P^k(x, \cdot), k, a/2)$. Now it must be demonstrated that there is but a single ergodic set with no cyclically moving subsets. If there were two ergodic sets, they would be invariant and disjoint, implying that $\alpha(P^k) = 0$ for all k. If there were cyclically moving subsets, $P^k(x, \cdot)$ would be concentrated on a particular one, depending on which subset contained x. This would imply that $\alpha(P^k) = 0$ for all k. Thus condition D_0 must hold.

Now suppose X satisfies Condition D_0 with triple (φ, k, δ). Lemma 17 implies that for $n > n_0$,

$$|P^n(x, A) - P^n(y, A)| < 2\rho^n$$

for all choices of $x, y \in E$, $A \in \mathcal{E}$. The constant ρ is strictly less than 1, so $\alpha(P^k) > 0$, for all n such that $2\rho^n < 1$. ∎

It is not surprising to find that many results that have been found for Doeblin recurrent chains have also been found for chains satisfying the Ergodic Condition. The interested reader sould see Iosifescu and Theodorescu (1969) for a comprehensive study from this viewpoint.

2.2. GSMPS AND RELATED PROCESSES

AMONG THE MOST interesting efforts to use stochastic processes as models for complex phenomena are the generalized semi-Markov processes suggested by Matthes (1962). The construction of these processes that follows is based largely on the presentations of Whitt (1976) and Hordijk and Schassberger (1976).

Let S be a finite set called the state space and E be a finite set called the event space. With each state $s \in S$, associate a positive integer n_s and an n_s-tuple of distinct elements of E,

$$E(s) = (e_1(s), e_2(s), \ldots, e_{n_s}(s)) \text{ with } e_i(s) \in E.$$

This event set lists the possible events which can occur in state s. With some abuse of notation, we shall write $e_i \in E(s)$ if e_i is a coordinate of the vector E(s). Also, with each $s \in S$, associate the space of clock readings

$$C(s) = \{ (c_1, c_2, \ldots, c_{|E|}) : c_i = 0 \text{ if } e_i \notin E(s), c_i > 0 \text{ otherwise} \},$$

where |E| is defined to be the number of elements in E. Each element of C(s) has a coordinate corresponding to each event in E. By convention, if the event e_i is not in E(s), the clock corresponding to e_i (c_i) will read 0 in each element of C(s). If $e_i \in E(s)$, then c_i records the amount of time remaining until the e_i occurs. The clock c_i and event e_i are said to be active in state s if $e_i \in$ E(s). Let G(s)$=\{ s \} \times C(s)$ and $\Omega' = \bigcup_{s \in S} G(s)$. Let $\mathcal{F}'$ be the σ-algebra generated by sets of the form (s,A), where $s \in S$ and A is a Borel set in $R^{|E(s)|}$ projected to the active components of C(s). Then $(\Omega', \mathcal{F}')$ is a measurable space. Denote by $(\Omega, \mathcal{F})$ the product space $((\Omega')^\infty, (\mathcal{F}')^\infty)$ and call $Z_n = (X_n, C_n)$ the coordinate mappings of Ω. The i^{th} coordinate of the vector C_n will be denoted $C_{n,i}$. An element of Ω' shall be called a Z-state to distinguish it from an element of S. Each point of Ω', therefore, contains an element of S and a vector of clocks indicating the time remaining until each scheduled event occurs.

To determine how $Z = (X, C) = ((X_n, C_n), n = 0, 1, \ldots)$ behaves, three families of functions are required. First consider a family of speeds $(k_{s,e}; s \in S, e \in E(s))$ with the properties:

(1) $k_{se} \geq 0$ for all $s \in S, e \in E(s)$;

(2) $\sum_{e \in E} k_{se} > 0$ for all $s \in S$;

(3) $\sum_{s \in S} k_{se} > 0$ for all $e \in E$.

Intuitively, k_{se} is the amount the clock associated with event e decreases in each unit of time when the system is in state s. The most natural value of k_{se} is 1. This means that the amount of time remaining until event e occurs decreases by one unit in each unit of real time that elapses. Occasionally, however, models require speeds that are not unity. In computer models, for example, processor sharing is sometimes viewed as a server performing in each unit of time 1/n units of service for each of the n jobs requiring his service. For this situation positive speeds other than unity are useful and state-dependent speeds are required.

When services are interrupted and resumed at a later time, speeds of zero are needed. An example of a model with this feature will be introduced in Section 3.2. A speed of zero is a technically tricky addition, however, and we will need to make an assumption to simplify the situation. Assumption (9) will formally state the requirement we need.

Now a mapping to govern how X changes when events occur must be defined. Let the family of mappings

$$p(s, e, s'): S \times E \times S \to [0, 1]$$

indicate the probability that s' is entered from s when event e occurs. Assume that

(4) if $p(s, e, s') > 0$ then $e \in E(s)$;

(5) if $p(s, e, s') > 0$ then $E(s) - \{e\} \subseteq E(s')$;

(6) if $p(s, e, s') > 0$ then $k_{se} > 0$;

(7) for each pair (s,e), $s \in S$, $e \in E(s)$, $\sum_{s' \in S} p(s, e, s') = 1$ when $k_{se} > 0$;

(8) for each pair $(s, s') \in S \times S$, there is a finite sequence of events and states $e_1, s_1, e_2, \ldots, e_n$, which satisfies

$$p(s_1, e_1, s_2)p(s_2, e_2, s_3)\ldots p(s_{n-1}, e_n, s') > 0.$$

(9) if $k_{se} = 0$ for a pair $s \in S$, $e \in E(s)$, there is a path s, $e_1, s_1, \ldots, s_{n-1}, e_n, s_n$ with the properties that $k_{s_n,e} > 0$ and $k_{s_i e_j} = 0$ for $i = 1, \ldots, n$ and $e_j \in \{ e \in E(s) \mid k_{se} = 0 \}$.

Property (5) means that we will assume a 'noninterruptive' scheme (Schassberger's terminology). If an event is active in a state and does not occur then it must be active in the next state.

When property (8) is satisfied the process is said to be irreducible. The sequence $(s, e_1, s_1, \ldots, e_n, s')$ will be called the path between s and s'.

Condition (9) is the technical convenience mentioned above. When a speed associated with an event e is zero in a particular $s \in S$, this condition requires that there is a path to a state s_n where the speed $k_{s_n e}$ is positive, while the other clocks with zero speeds in state s also have a zero speed in *each* state of the path. This means the zero speeds can be made positive one at a time. This property will be convenient in our discussion of models with zero speeds in Section 3.1.

The scheduling of events is governed by a family of lifetime distributions $\{F(\cdot, s, e, t, e'); s \in S, e \in E(s), t > 0, e' \in E\}$. The distribution $F(x, s, e, t, e')$ represents the probability that the clock associated with event e will be set at a value less than or equal to x when there is a transition into state s after the occurence of event e' with the clock associated with event e reading t at the instant of transition. For each $s \in S$, $e' \in E$, $e \in E(s)$, these distributions must possess the following properties:

(10) for $t = 0$, $F(0,s,e,t,e') = 0$;

(11) for $t > 0$, $F(x,s,e,t,e') = 1_{[t,\infty]}(x)$;

(12) for $t = 0$, have densities with respect to Lebesgue measure which are bounded by K and have support on $[0, a_{s,e,e'})$, $a_{s,e,e'} \leq +\infty$. Also suppose that the distributions have finite first moments. Throughout this study, $Y_{s,e,e'}$ will be a random variable with distribution $F(\cdot, s, e, 0, e')$.

Assumption (11) implies that if a clock is active at the time of the occurrence of an event it is not reset or become inactive when that event occurs, but reads the same immediately after the event as it did immediately before the event.

For a pair $(s,c) \in \Omega'$ let

$$t^* = t^*(s,c) = \min \frac{c_i}{k_{s,e_i}}$$

where the minimization is taken over those coordinates with $c_i k_{s,e_i} > 0$. Intuitively, t^* is the minimum time until one of the active clocks of s hits zero. Let $e^* = e^*(s,c)$ be the event associated with the clock reading t^*. The variables t_i^* and e_i^* will denote the minimum clock reading and its associated event for the i^{th} coordinate of Ω.

A probability transition function can now be defined for the process Z. For a set

$$A = \{ s' \} \times \{ \times_{j=1}^{|E|} [0, \ell_j) \}$$

with $\ell_j = 0$ if $e_j \notin E(s)$, $\ell_j > 0$ otherwise; define $P: \Omega' \times \mathcal{F}' \to [0,1]$ by

$$P((s,c), A) = p(s, e^*(s,c), s') \prod_{j=1}^{|E(s')|} F(\ell_j, s', e_j, t_j, e^*) \tag{13}$$

where

$$t_j = \begin{cases} 0, & \text{if } e_j \notin E(s), \\ c_j - k_{s,e_j} t^*, & \text{if } e_j \in E(s). \end{cases}$$

Let $\{ \mathrm{P}((s,c),\cdot) : s \in S, c \in C(s) \}$ be the family of distributions that govern transitions in the Z chain. If $F(x,s,e,t,c')$ is a measurable function of t for each $x \in [0,+\infty), s \in S, e \in E(s), c' \in E$, then P is a Markov kernel (see Definition 2.1.1 or Revuz, 1975). Finally,, let ν be a measure on $(\Omega', \mathcal{F}')$ and call it the starting measure if $\mathrm{P}(Z_0 \in A) = \nu(A)$.

(14) DEFINITION. A process $Z = \{ Z_n, n \geq 0 \}$ constructed in this way is called a Generalized Semi-Markov Ordered Pair (GSMOP).

A continuous time process related to the GSMOP is the process of the most interest; it will be constructed from the GSMOP. First let

$$Q(n) = \sum_{m=0}^{n-1} \min_{C_{m,i} \cdot k_{X_{m},e_i} > 0} \frac{C_{m,i}}{k_{X_{m},e_i}}$$

and

$$N(t) = \max\{ n > 0 : Q(n) \leq t \}.$$

The processes represent the time of the n^{th} transition and the number of transitions by time t, respectively.

(15) DEFINITION. The process $\mathfrak{Z} = \{ \mathfrak{Z}_t, t \geq 0 \} = \{ (\mathfrak{X}_t, \mathfrak{C}_t), t \geq 0 \}$ where

$$\begin{aligned} \mathfrak{Z}_t = (\mathfrak{X}_t, \mathfrak{C}_t) = (X_{N(t)}, C_{N(t),1} - k_{X_{N(t)},e_1}(t - Q(N(t))), \ldots, \\ C_{N(t),|E|} - k_{X_{N(t)},e_{|E|}}(t - Q(N(t))). \end{aligned}$$

is called a Generalized Semi-Markov Order (GSMO). Its first component is the generalized Semi-Markov Process (GSMP).

The only standard terminology here is the GSMP — each author has his own terminology for the other processes. It is important to note that the GSMOP is a Markov chain on a general state space and the GSMO is a Markov process on the same state space.

An example will help clarify the notation.

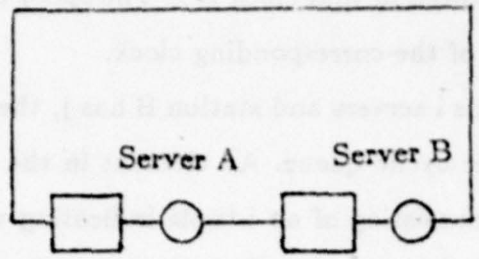

Figure 1. Two Station Single Server Cyclic Queue

(16) EXAMPLE. Two station single server cyclic queue with k customers.

This system is an arrangement of two simple queues with the departure process of each queue the arrrival process of the other. Figure 1 shows a schematic diagram of this system. In this figure (as in all others in this report) a circle represents a server while a rectangle represents a waiting room.

The state space S for the queue length process of this system could be the collection of ordered pairs $((0,k),(1,k-1),\ldots,(k,0))$ where each pair $(i,k-i)$ indicates that i customers are waiting for or are being served by server A while k-i are doing the same for B. The only events that can occur are a service completion by A or one by B, implying that the clock vector c is a pair. The first component is active in states $((1,k-1),\ldots,(k,0))$ while the second is active in $((0,k),\ldots,(k-1,1))$. Suppose the initial distribution of the process is $\nu(A) = P\{((0,k),(0,c)),A\}$. Note that the choice of c is immaterial so long as it is positive.

Figure 2 illustrates a sample path of GSMO and its related processes. The 4-tuple (number of customers at station A, number of customers at B, time remaining in service at station A, time remaining in service at station B) for $t \geq 0$ is a GSMO. The GSMOP is the collection of readings of the GSMO at the times events occur (dotted lines). The first component of the GSMO is a GSMP. (In this case, it is the order pair describing the number of jobs waiting at each server.) Notice that the slope of each clock is -1. This

indicates that the speed of that clock is 1. The speed of service is the absolute value of the slope of the corresponding clock.

If station A has i servers and station B has j, the system will be referred to as an (i,j)-server cyclic queue. An element in the state space S would be an ordered triple consisting of an i-tuple indicating which servers are active at station A, a j-tuple measuring the same thing at station B, and an integer between 0 and k representing the number of jobs waiting and in service at station A. There would be i+j clocks in the event space — one for each server. If, after completion of service at station A, a job may reenter the waiting room of A with positive probability, the system will be said to have feedback.

Throughout this report, the limiting behavior of GSMOs and GSMOPs is a question of great interest and a note about the existence of invariant measures should be made here. Results in Sections 3.1 and 4.2 allow the classification of GSMOPs by recurrence properties and the application of results in Section 2.1. Invariant measures exist for all GSMOPs which satisfy the assumptions (1)-(12) and the additional stipulation that $a_{s,e,e'}$ is either finite or infinite for all $s \in S$, $e \in E$, $e' \in E(s)$. Whitt (1976) demonstrates the existence of such a measure under a slightly different construction. The proof can be easily modified to the situation specified here. We shall indicate the limiting distribution associated with a process by a prime. For example, the limiting random variable of Z is Z', and $\mathfrak{Z}$ is $\mathfrak{Z}'$.

Finally, occasionally more than one GSMOP or GSMO will be under discussion. When there is danger of confusion, a subscript will be appended to the notation to indicate which process is concerned, as in Ω'_W or P_Z.

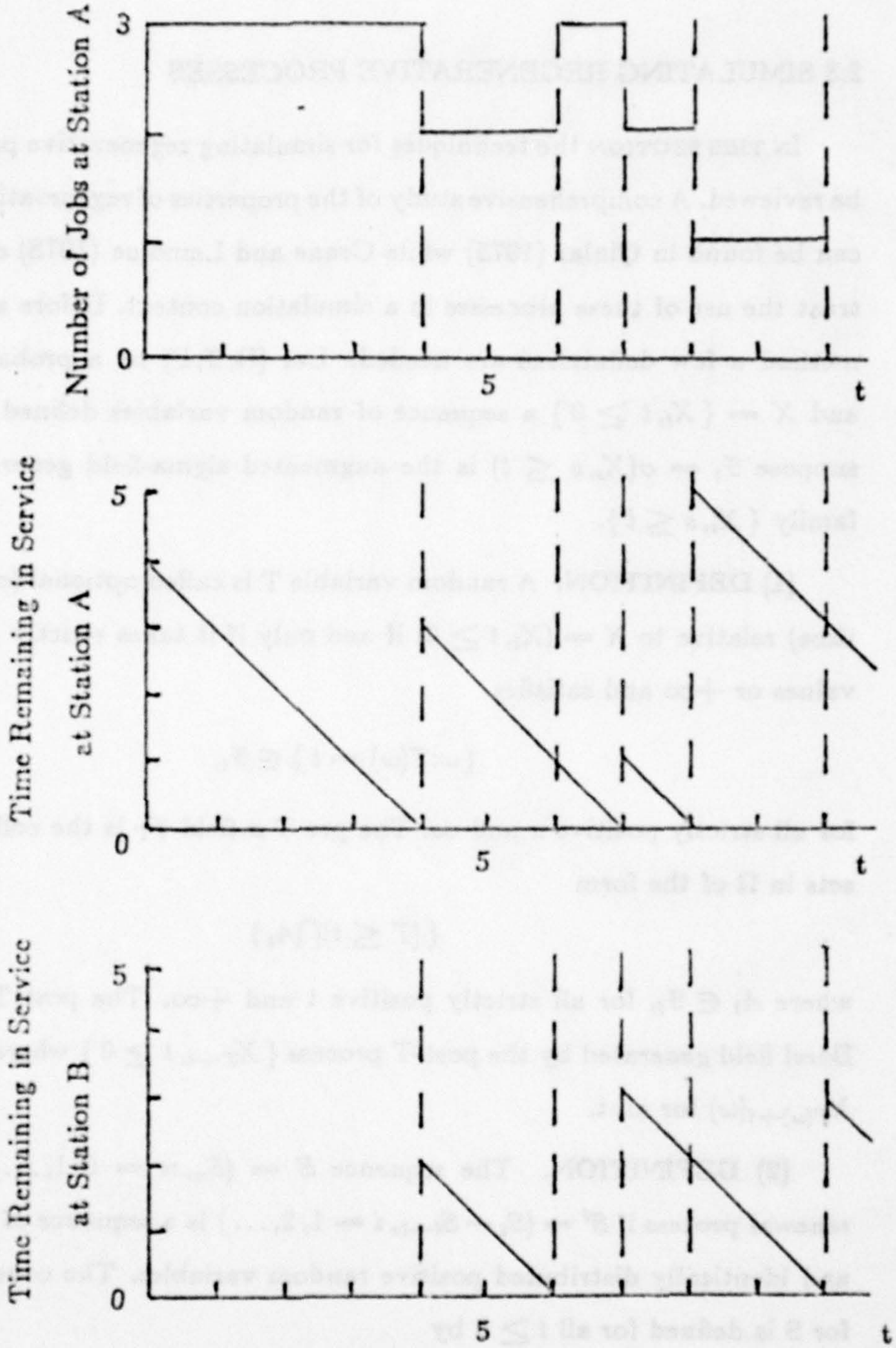

Figure 2. A GSMO Sample Path

2.3 SIMULATING REGENERATIVE PROCESSES

IN THIS SECTION the techniques for simulating regenerative processes will be reviewed. A comprehensive study of the properties of regenerative processes can be found in Çinlar (1975) while Crane and Lemoine (1978) exhaustively treat the use of these processes in a simulation context. Before studying the method a few definitions are needed. Let $(\Omega, \mathfrak{F}, P)$ be a probability space and $X = \{ X_t, t \geq 0 \}$ a sequence of random variables defined on Ω. Also suppose $\mathfrak{F}_t = \sigma(X_s, s \leq t)$ is the augmented sigma-field generated by the family $\{ X_s, s \leq t \}$.

(1) DEFINITION. A random variable T is called optional (or a stopping time) relative to $X = (X_t, t \geq 0)$ if and only if it takes strictly positive real values or $+\infty$ and satisfies

$$\{ \omega : T(\omega) = t \} \in \mathfrak{F}_t,$$

for all strictly positive n and ∞. The pre-T σ-field $\mathfrak{F}_T$ is the collection of all sets in Ω of the form

$$\{ (T \leq t) \bigcap A_t \}$$

where $A_t \in \mathfrak{F}_t$, for all strictly positive t and $+\infty$. The post-T field is the Borel field generated by the post-T process $\{ X_{T+t}, t \geq 0 \}$ where $X_{T+t}(\omega) = X_{T(\omega)+t}(\omega)$ for all t.

(2) DEFINITION. The sequence $S = (S_n, n = 0, 1, \ldots)$ is called a renewal process if $S' = (S_i - S_{i-1}, i = 1, 2, \ldots)$ is a sequence of independent and identically distributed positive random variables. The counting process for S is defined for all $t \geq 0$ by

$$N(t) = \sup(n : S_n \leq t).$$

An important result about the counting process is the Elementary Renewal Theorem.

(3) LEMMA. If the renewal process S satisfies the property that $S_n < +\infty$ almost surely for all n, then

$$\lim_{t \to \infty} \frac{E\{N(t)\}}{t} = \frac{1}{E\{S_n - S_{n-1}\}}$$

Proof. See Çinlar (1975). If S is integer valued, then the counting process N can also be integer valued and a corresponding result can be obtained for $\lim_{i \to \infty} E\{N_i\}/i$. ∎

(4) DEFINITION. The process $\{X_t, t \geq 0\}$ is said to be regenerative provided there exist a sequence $S_0, S_1, \ldots$ of stopping times such that

(a) $S = \{S_n, n \geq 0\}$ is a renewal process; and

(b) for any $n, m \in N, t_1, t_2, \ldots, t_n \in R^+$, and any bounded function f defined on Ω^n;

$$E\{f(X_{S_m+t_1}, \ldots, X_{S_m+t_n}) \mid X_t, t < S_m\} = E\{f(X_{t_1}, X_{t_2}, \ldots, X_{t_n})\}.$$

Intuitively, the second condition requires that the process after S_m be independent of the past and the probability law governing the post-S_m process be the same as thate governing the original process.

EXAMPLES

(5) Let $\{X_n, n \geq 0\}$ be a recurrent Markov chain on state space $(1, 2, \ldots, k)$, with initial state i. If $Z_t = X_n$ for all $t \in [n, n+1)$, then Z is a regenerative process with the n^{th} regeneration time being the n^{th} entrance to j.

(6) Let $\{Z_t, t \geq 0\}$ be the queue length process of a GI/G/1 queue, with traffic intensity $\rho < 1$. The regenerative times are the times of these arrivals which find an empty system.

The idea behind the regenerative method is the division of a sample path of the process into pieces which are independent and identically distributed, and the formation of a central limit theorem based on those i.i.d. pieces. Let $\{X_t, t \geq 0\}$ be a regenerative process on R with a sequence of regeneration times $\{\beta_n, n \geq 0\}$. Let F be the common distribution of $\alpha_n = \beta_n - \beta_{n-1}$, and suppose $E\{\alpha_n\} < +\infty$. If F is not arithmetic (see Çinlar (1975)), it is known that there exists a random variable X such that $X_t \Rightarrow X$. Here $\Rightarrow$ denotes weak convergence or convergence in distribution ($P(X_t \leq x) \rightarrow P(X \leq x)$ as $t \rightarrow +\infty$ for all x for which $P(X \leq x)$ is continuous.) Let f be a measurable function from E to R and define a sequence $\{Y_k, k = 1, 2, \dots\}$ where

$$Y_k = \int_{\beta_{k-1}}^{\beta_k} f(X_s)\, ds$$

Three results are important in the simulation of regenerative processes.

(8) LEMMA. The sequence $\{(Y_k, \alpha_k), k = 1, 2, \dots\}$ is independent and identically distributed.

Proof. Obvious from the properties of regenerative processes. ∎

(9) LEMMA. If $E\{f(X)\} < +\infty$, then $r \equiv E\{f(X)\} = E\{Y_1\}/E\{\alpha_1\}$.

Proof. See Crane and Iglehart (1975). ∎

An application of the standard central limit theorem yields

(10) LEMMA. Let $Z_k = Y_k - E\{f(X)\}\alpha_k$, with $\sigma^2 = \operatorname{Var} Z_k < +\infty$. Then

$$\frac{\sum_{k=0}^{n} Z_k}{\sigma\sqrt{n}} \Rightarrow N(0, 1).$$

Proof. See Chung (1974). ∎

Typically the variance constant σ is unknown, and must be estimated. The classical estimate is $s^2 = s_{1,1} - 2\hat{r}s_{1,2} + \hat{r}^2 s_{2,2}$ where $\hat{r} = \overline{Y}/\overline{a}$ is the sample estimate of $E\{fX\}$ and $s_{1,1}$, $s_{2,2}$ and $s_{1,2}$ are the sample variances of $\{Y_n, n \geq 1\}$, $\{a_n, n \geq 1\}$ and their sample covariance. It can be shown that $\hat{r} \rightarrow r$ and $s^2 \rightarrow \sigma^2$ with probability 1 as the sample size increase to $+\infty$. A continuous mapping argument is all that is required in order to substitute s for σ in (10).

CHAPTER III

GSMOS WITH SINGLE SETS

3.1. IDENTIFYING REGENERATIVE GSMOS

THE PURPOSE OF this section is to find conditions for a GSMO to be regenerative in the special case where $a_{s,c,c'} = +\infty$ for all s, c and c'. We seek a sequence of stopping times where the future of the process at each time is governed by the same probability distributions. It would be convenient if each stopping time in the sequence is also the time an event occurs. This would mean that examining the transition function of the associated GSMOP should tell us some information about the stopping times. In a GSMOP, a particular Z-state may not be visited infinitely often, since the clocks have positive densities on $[0, \infty)$. Therefore, we must find a set $A \subseteq \Omega'$ where $\mathrm{P}((x,c),B) = \mathrm{P}((y,d),B)$ for all (x,c), $(y,d) \in A$, $B \in \mathcal{F}'$. (Recall that $\mathcal{F}'$ is the σ-algebra we constructed on Ω'.)

Note that if there is a clock active at the time of a transition, the transition function P defined in 2.2.13 guarantees that it is not reset at the time of any event. This means that the reading of that clock determines a subset of Ω' to which the next Z-state is confined. (Namely, that subset which has precisely the same readings for the active clocks.) Different clock readings therefore

determine different subsets of Ω' to which the next Z-state is confined. The probability that any clock has the same reading at an infinite number of transitions is 0 since the clock's density is positive on $[0,\infty)$. This means that the event that governs a transition from a Z-state in A must be the only event active in that Z-state. This condition is formalized in the following definition.

(1) DEFINITION. A GSMOP, GSMO or GSMP will be called single if there exists a nonempty set $S_1 \subseteq S$ with $|E(s)| = 1$ for all $s \in S_1$. A state $s \in S_1$ will be called a single state and the set G(s) a single set.

EXAMPLES.

(2) (3,1) server cyclic queues with k jobs and feedback. When all the jobs in the system are waiting for the single server (state ((0,0,0),1,0) in the state space described in Example 2.2.18) the only active clock is the one which measures the remaining time in service for that server.

(3) (2,2) server cyclic queues with k customers. As long as the number of jobs in the system exceeds 2, there will be at least 2 servers active at all times and a single set does not exist. A closed network must allow all customers to wait for a single server in order for a single set to exist.

(4) A single server simple queue with traffic intensity less than one. The state which represents an empty system is the single state. The only clock active in this state is one which measures the amount of time remaining until the next arrival. If an open network has a single input stream when the system is empty, then a single set exists.

An interesting observation is that all GSMOPs with single sets do not satisfy Doeblin's condition. (My thanks go to Phil Heidelberger for pointing out the principle underlying this example.) Example (2) does not for $k > 4$. Consider, for a fixed positive c, the set

$$A(c) = \{(((1,\delta,1),\gamma,j),c_1,c_2,c_3,c_4): j = 2,\ldots,k;$$
$$\delta = 0,1; \gamma = 0,1; c_1 - c_3 = c > 0, c_1, c_4 \in R^+\}$$

This set is the collection of Z-states which have the first and third servers at station A active, with the difference in their remaining services equal to c. For any $n \geq 1$, $\epsilon > 0$, and any choice of c_2, c_4 and c, c_1 and c_3 can be chosen so that

$$(((1,1,1),1,k-1)c_1,c_2,c_3,c_4) \in A(c)$$

and

$$\mathrm{P}^n\{(((1,1,1),1,k-1),c_1,c_2,c_3,c_4),A(c)\} > 1-\epsilon.$$

In words this means that when 2 of the servers at station A are busy with a difference of c between the times remaining in their respective services, we can choose the times until those servers complete service to be so enormous that even after a large number of transitions the difference between the clocks is still c with a high probability, since the probability that either server has completed service is small. There are also choices of c_1 and c_3 which satisfy similar conditions for A(d). As long as $c \neq d$, this implies that the coefficient of ergodicity is zero. There is only one ergodic set which has no cyclically moving subsets, so, by Theorem 2.1.19, Doeblin's Condition is not satisfied. It is easy to see that single GSMOPs satisfy condition (ii) of Definition 2.1.4(b) for $(G(s),\Omega',P((s,c),\cdot),1,1)$-recurrence if $s \in A$, $(s,c) \in G(s)$. Is condition (i) of the definition satisfied? The following theorem is helpful in answering this question.

(5) THEOREM. Let Z be a GSMOP with single state y. Suppose for some $s_1 \in S$, the path to y is $(s_1,e_1,s_2,\ldots,s_{n-1},e_n,y)$. Of the events in this path, let the set $\{e_{i_1},e_{i_2},\ldots,e_{i_k}\}$, $1 = i_1 < i_2 < \ldots < i_k$, be elements of

E(s_1). Define, for $c \in C(s_1)$, and $l = 1, 2, \ldots, |E(s_1)|$,

$$
\begin{aligned}
t'_1 &= \frac{c_{i_1}}{k_{s_1 e_1}}, \\
t'_l &= \frac{c_{i_l} - \sum_{m=1}^{l-1} k_{s_{i_m} e_{i_{m+1}}} t'_m}{k_{s_{i_l} e_{i_l}}},
\end{aligned}
\tag{6}
$$

Let $\Upsilon(s_1)$ be the subset of C(s_1) where

$$
\begin{aligned}
&\frac{c_{i_1}}{k_{s_0 e_{i_1}}} < \frac{c_{i_j}}{k_{s_0 e_{i_j}}} \quad \text{for } j = 2, \ldots, k, \\
&\frac{c_{i_l} - \sum_{m=1}^{l-1} k_{s_{i_m} e_{i_l}} t'_m}{k_{s_{i_{l-1}} e_{i_l}}} < \frac{c_{i_j} - \sum_{m=1}^{l-1} k_{s_{i_m} e_{i_j}} t'_m}{k_{s_{i_{l-1}} e_{i_j}}},
\end{aligned}
\tag{7}
$$

for $j > l, l = 1, 2, 3, \ldots, |E(s_1)|$, is satisfied. Then $\Upsilon(s_1)$ is not empty. Also, if $c \in \Upsilon(s_1)$, then $P^n((s_1, c), G(y)) > 0$.

Note: (a) In this theorem, k is either $|E(s_1)|$ or $|E(s_1)| - 1$.

(b) In equation (6), $\sum_{i=1}^{j} t'_i$ is the time of the j^{th} event if the events that are active in s_1 are allowed to occur consecutively in the order prescribed by $i_1, i_2, \ldots, i_k$. These times are necessary for condition (7) which requires that the clocks active in s_1 be in the right order for the path from s_1 to y to be followed.

(c) The idea of this proof is that if the active clocks of state s_1 are in the proper order for the path to y to be followed, the other events that must be scheduled along the path can be scheduled with positive probability on a set that does not disturb the order of events prescribed in (7). Furthermore the events not active in s_1 but activated in the path from s_1 to y can be scheduled in such a way that the path is followed with postive probability.

Proof. The assertion is trivially true for n=1. Now suppose the assertion is true for $n = 0, 1, ..., m-1$. Now let $n = m$. To see the set $\Upsilon(s_1)$ is nonempty, choose $c_2 \in \Upsilon(s_2)$ and set $c_{1i} = 0$ if $e_i \notin E(s_1)$. To the remaining positive elements of c_2 add $(\epsilon k_{s_1 e_i})/k_{s_1 e_{i_1}}$, for some $\epsilon > 0$, and set $c_{1e_i} = c_{2e_i} + \epsilon k_{s_1 e_i}/k_{s_1 e_i})$

Let (resetting it if necessary) the entry for $c_{1e_{i_1}}$ equal $\epsilon/k_{s_1e_{i_1}}$. This new clock vector must be an element of $\Upsilon(s_1)$.

Note that

$$\begin{aligned}
P^m((s_1,c_1),G(y)) &\geq \int_{G(s_2)} p(s_1,e_1,s_2)P^{m-1}((s_2,c_2),G(y))\,P((s_1,c_1),d(s_2,c_2)),\\
&\geq \int_{\Upsilon(s_2)} p(s_1,e_1,s_2)P^{m-1}((s_2,c_2),G(y))\,P((s_1,c_1),d(s_2,c_2)).
\end{aligned}$$

If $E(s_2) \subseteq E(s_1) - \{e_{i_1}\}$, then $P((s_1,c_1),\Upsilon(s_2)) = 1$ and the assertion is true by the inductive hypothesis. If $E(s_2) \not\subseteq E(s_1) - \{e_{i_1}\}$, the clocks activated in state s_2 must also satisfy not disturb the sense of the inequalities in (7) with positive probability. Since $\Upsilon(s_2)$ is nonempty and the relations in (7) are strict inequalities, $\Upsilon(s_2)$ must contain a rectangle in the coordinates of $E(s_2) - \{E(s_1) - \{e_{i_1}\}\}$ which does not disturb the order of the clocks prescribed in (7). The clocks associated with the event must assign positive mass to the appropriate side of the rectangle, since they have densities that are positive on the entire real line, so (7) and the inductive hypothesis imply the desired conclusion. ∎

This allows us to show that GSMOPs with single sets are Harris recurrent chains.

(8) THEOREM. Let Z be a GSMOP with single state y. Then Z is $(G(y),\Omega',P((y,c),\cdot),1,1)$-recurrent for any $c \in C(y)$.

Proof. For any $c \in C(y)$ Condition (ii) of Definition 2.1.4(b) is trivially satisfied by $(G(y),\Omega',P((y,c),\cdot),1,1)$ since $P((y,c),A) = P((y,d),A)$ for all $c,d \in C(y)$, $A \in \mathcal{F}'$. It remains to be shown that $P_z(Z_n \in G(y)$ for some n$) = 1$ for all $z \in \Omega'$. Since S has only a finite number of elements ($X_n = s$ i.o.) for some $s \in S$. Suppose that for each $c \in C(s)$ there is a finite sequence of measurable sets $B_1, B_2, \ldots, B_{u(s,c)}$ which satisfies

$$P_{(s,c)}\{Z_1 \in B_1, Z_2 \in B_2, \ldots, Z_{u(s,c)} \in B_{u(s,c)}, Z_{u(s,c)+1} \in G(y)\} > 0. \quad (9)$$

Now define

$$T(1) = \min(n: X_n = s),$$
$$T(n) = \min(n > T(n-1) + u(s, C_{T(i-1)}): X_n = s).$$

It is clear that

$$\mathrm{P}_{(s,c)}(Z_{T(i)+u(s,c_{T(i-1)})} \in (G(y))^c \text{ for all i}) = 0.$$

This would complete the proof of the theorem provided such a sequence of B sets can be constructed for any $c \in C(s)$. To find the sequence of sets for a given $c \in C(s)$, first suppose all the speeds of the process are positive. The condition is satisfied by choosing a path from s to y in two pieces. First select a path that could be followed if all the events scheduled in s occurred consecutively. Say the path this generates is $s, e_1, s_1, \ldots, e_{|E(s)|}, s'$. This means that the clocks for events activated in $e_1, \ldots, e_{|E(s)|}$ must be set at a large enough value so that newly activated clocks occur after all the events active in s. This occurs with positive probability since all the clocks have positive densities on the entire positive half line. To find the remainder of the path from s to y, choose the path between s' and y that exists by the irreducibility assumption. The clock vector c satisfies (7) for this path, so Theorem 5 then implies the desired conclusion.

Speeds of zero are slightly more troublesome. The above path utilizing the events of s may not be possible if the speed of an event in E(s) is zero for every state in every path of |E(s)| steps from s. This means that it may not be possible for us to construct the first piece of the path as we did above. There must, however, be a path from s to s' in S, say $s, e_1, s_1, \ldots, e_j, s'$ where

(a) $e_i \in E(s)$ for $i = 1, \ldots, j$;

(b) $k_{s'e} = 0$ for all $e \in \Delta = E(s) - \{e_1, \ldots, e_j\}$.

This path may be obtained by examining all the paths of length $|E(s)|$ from s and choosing the path which has the most elements in E(s) occurring

consecutively and which satisfies the appropriate subset of inequalities given in (7). To satisfy the remaining inequalities in (7) the process should visit states where the events in (δ) have positive clocks. In particular, it should visit those states guaranteed in (2.2.9) where the speeds of events in Δ are made positive one at a time. The irreducibility assumption guarantees that one state for each event in Δ can be visited. The inequalities in (7) are fulfilled since only one event at a time is assigned a positive speed. When all these sets have been visited, the irreducibility assumption implies that G(y) can be reached. By construction Theorem 5 then implies the desired result. ∎

It is not clear that assumption 2.2.9 is necessary to insure the validity *of this result. If the assumption is not made, more than one clock can switch* from a zero speed to a positive one after an event occurs and it becomes difficult to determine which of these newly activated clocks expires first.

The possibility of a clock speed of zero is responsible for the complexity of this proof. Were it not for positive densities on the entire real line, the result might not be true. When all speeds are unity a cleaner proof is possible (see Section 4.2).

This result is one way to establish that the chain will visit the single set infinitely often. The reasoning in the theorem can also be used to determine that the expected number of steps between successive visits to the single set is finite. Combined with the fact that the process leaves the single set the same way each time, the key ingredients for regenerative processes are present. Visits to the single set break the process into cycles which are independent and *identically distributed random variables. The formalities of demonstrating* this fact begins with the definition of the shift operator.

(10) DEFINITION. Let $(E,\mathcal{E})$ be a measurable space and let $E^{\infty} = E \times$

$E \times \ldots$. The shift θ is a mapping of E^∞ such that

$$\theta{:}\omega = \{\omega_n \in E, n \in N\} \mapsto \theta\omega = \{\omega_{n+1}, n \in N\}.$$

Define its iterates by composition: θ^0 is the identity and $\theta^k = \theta \circ \theta^{k-1}$ for $k \geq 1$. Each of these shift operators induces an inverse set mapping. These are defined by

$$\theta^{-n}(A) = \{\omega \in E \mid \theta^n(\omega) \in A\}.$$

A random or α-shift can be defined by

$$\theta^\alpha(\omega) = \theta^n(\omega) \quad \text{on } \{\omega{:}\alpha(\omega) = n\}.$$

The inverse α-shift is, of course, $\alpha^{-1}(A) = \{\omega{:}\alpha(\omega) \in A\}$.

Suppose that Z is a $(G(s), \Omega', P((s,c),\cdot),1,1)$-recurrent Markov chain. Then the successive entrance times to the single set are optional relative to $\{Z_n, n \geq 0\}$ and are defined by:

$$\begin{aligned} \beta_1 &= \min(n > 0 \mid Z_n \in G(s)), & \alpha_1 &= \beta_1; \\ \beta_i &= \min(n > \beta_{i-1} \mid Z_n \in G(s)), & \alpha_i &= \beta_i - \beta_{i-1}. \end{aligned}$$

Let the pre-β_i and post-β_i σ-fields be denoted by $\mathcal{F}_{\beta_i}$ and $\mathcal{F}'_{\beta_i}$, respectively. For convenience, let the initial distribution satisfy

$$\nu(A) = P((s,c), A)$$

for $(s,c) \in G(s)$, $A \in \mathcal{F}'$.

(11) THEOREM. Let Z be a GSMOP with single state s and single set G(s). Then, for each i, the Borel fields $\mathcal{F}_{\beta_i}$ and $\mathcal{F}'_{\beta_i}$ are independent. Also the post-β_i process has the same distribution as the original process.

Proof. Suppose $\Lambda \in \mathcal{F}_{\beta_i}$ and $B_j \in \mathcal{F}'_{\beta_i}$, $1 \leq j \leq k$, $k \geq 1$. The result will be established if the equality

$$P_{Z_0}(\Lambda, Z_{\beta_i + j} \in B_j, 1 \leq j \leq k) = P_{Z_0}(\Lambda) P_{Z_0}(Z_j \in B_j, 1 \leq j \leq k).$$

is demonstrated. The expression $\Lambda \bigcap \{ \beta_i = n \}$ can be rewritten as $\Lambda_n \bigcap \{ \beta_i = n \} = \sigma(Z_i, i \le n)$ where $\Lambda_n \in \mathcal{F}_n$ for each n. Therefore,

$$\begin{aligned}
&P_{Z_0}(\Lambda, Z_{\beta_i+j} \in B_j, 1 \le j \le k) \\
&\qquad = \sum_{n=0}^{\infty} P_{Z_0}(\Lambda_n, \beta_i = n, Z_{\beta_i+j} \in B_j, 1 \le j \le k) \\
&\qquad = \sum_{n=0}^{\infty} P_{Z_0}(\Lambda_n, \beta_i = n) P_{Z_0}(Z_j \in B_j, 1 \le j \le k) \\
&\qquad = P_{Z_0}(\Lambda) P_{Z_0}(Z_j \in B_j, 1 \le j \le k)
\end{aligned}$$

The second equivalence is justified by the fact the Z is a Markov chain. ∎

(12) COROLLARY. For $G \in \mathcal{F}'$, $P_{Z_0}(\theta^{-\beta_i}(G)) = P_{Z_0}(G)$.

Proof. Define the set function P′ by

$$P'(B) = P_{Z_0}(\theta^{-\beta_i}(B)).$$

The negative shift operator $\theta^{-\beta_i}$ maps disjoint sets to disjoint sets, so P′ can easily be shown to be a probability measure. If G is a finite product set, such as

$$G = \bigcap_{j=1}^{k} \{ z \mid z_{n_j} \in B_{n_j} \}$$

where $\{ n_1, n_2, \ldots, n_k \}$ is an arbitrary finite subset of the positive integers B_{n_j} is a rectangle in Ω', then it follows from the Theorem that P and P′ agree on sets of this form and thus on their finite unions and complements. The measure P is σ-finite, so the same is true for every $G \in \mathcal{F}'$ (see Chung, 1974, Theorem 2.2.3). ∎

(13) THEOREM. The random vectors $\{ V_k, k = 1, 2, \ldots \}$ are independent and identically distributed, where $V_k = \{ \beta_i, Z_{\beta_{i-1}+1}, \ldots, Z_{\beta_i} \}$.

Proof. To show the vectors are identically distributed, without loss of generality V_1 and V_2 can be examined. For $n \in N$ and an n-dimensional set $B \in \mathcal{F}_n$, it

is true that

$$\{z \in \Omega : \alpha_2(z) = n, (Z_{\alpha_1+1}(z), \ldots, Z_{\alpha_1+\alpha_2}(z)) \in B\}$$
$$= \{z : \alpha_1(\theta^{\alpha_1}(z)) = n, (Z_1(\theta^{\alpha_1}(z), \ldots, Z_{\alpha_1}(\theta^{\alpha_1}(z))) \in B\}$$
$$= \{z' = \theta^{\alpha_1}(z) : \alpha_1(z') = n, (Z_1(z'), \ldots, Z_{\alpha_1}(z')) \in B\}$$
$$= \theta^{-\alpha_1}\{z : \alpha_1 = n, (Z_1(z), \ldots, Z_n(z)) \in B\}.$$

The preceding corollary proves that this set has the same probability as $\{z : \alpha_1 = n, (Z_1(z), \ldots, Z_{\alpha_1}(z)) \in B\}$. Since $\alpha^k = \alpha \circ \theta^{\beta_{k-1}}$, it is clear that $V_k \in \mathcal{F}'_{\beta_{k-1}}$ while $V_1, \ldots, V_{k-1} \in \mathcal{F}_{\beta_{k-1}}$, so Theorem 11 establishes their independence. ∎

(14) COROLLARY. The random variables $\{Y_k, k = 1, 2, \ldots\}$ are independent and identically distributed where

$$Y_k = \sum_{n=\beta_{k-1}+1}^{\beta_k} f(Z_n)$$

and f is measurable on Ω'.

Proof. If $g_n(Z_1, \ldots, Z_n) = \sum_{i=1}^{n} f(Z_i)$ then g_n is measurable. For an open set A in R, let $B = g_n^{-1}(A)$ in Theorem 13 and the result is immediate. ∎

A central limit theorem for a GSMOP with a single set is the next order of business. Consider the random variables $\{Y_k - (EY/E\alpha)\,\alpha_k : k \geq 1\}$ where EY and $E\alpha$ are the common means of their respective sequences. They are independent and identically distributed with mean 0 and variance σ^2. If σ^2 is finite, the standard central limit theorem for this sequence is

$$\frac{\sum_{k=1}^{n} [Y_k - (EY/E\alpha)\,\alpha_k]}{\sigma\sqrt{n}} \Rightarrow N(0,1).$$

The confidence intervals desired are for a function of Z', so it would be desireable if $EY/E\alpha$ could be replaced by Ef/Z'. The next theorem demonstrates how this can be accomplished.

(15) **THEOREM.** Let Z be a GSMOP with a single set and a limiting random variable Z' distributed according to π. Let $f:\Omega' \mapsto R$ be measurable and suppose $\sigma^2 < +\infty$. Then

$$\frac{\sum_{k=1}^{n}(Y_k - \mathrm{E}\{f(Z')\}\alpha_k)}{\sigma\sqrt{n}} \Rightarrow N(0,1).$$

Proof. The definition of the sequence $\{\beta_i, i \geq 1\}$ insures that for each n>0 there exists a unique integer $l(n)$ such that $\beta_{l(n)} < n \leq \beta_{l(n)+1}$. The cumulative process S_n can be written as

$$S_n = \sum_{i=0}^{n} f(Z_i) = \sum_{j=1}^{l(n)} Y_j + \sum_{i=\beta_{l(n)}+1}^{n} f(Z_i).$$

The ratio limit theorem (Theorem 2.1.6) implies that $\lim_{n\to+\infty} \mathrm{E}\{S_n\}/n \to \mathrm{E}\{f(Z')\}$. The theorem then depends on the validity of the statement

$$\lim_{n\to+\infty} \mathrm{E}\{S_n\}/n \to \mathrm{E}\{Y\}/E\{\alpha\}.$$

First note that

$$\left|\lim_{n\to+\infty} \mathrm{E}\left\{\sum_{i=\beta_{l(n)}+1}^{n} f(Z_i)\right\}\right| \leq \lim_{n\to+\infty} \mathrm{E}\left\{\left|\sum_{i=\beta_{l(n)}+1}^{n} f(Z_i)\right|\right\}$$

$$\leq \lim_{n\to+\infty} \mathrm{E}\left\{\sum_{\beta_{l(n)}+1}^{\beta_{l(n)+1}} |f(Z_i)|\right\}$$

$$\leq \lim_{n\to+\infty} \mathrm{E}\left\{\max_{0\leq k\leq n} \sum_{\beta_{k-1}+1}^{\beta_k} |f(Z_i)|\right\}$$

According to Corollary 4, the random sums $\{Y_k : k \geq 0\}$ above are independent and identically distributed. The assumption that $\sigma^2 < +\infty$ implies that $E\{\sum_{i=\beta_{k-1}+1}^{\beta_k} |f(Z_i)|\} < +\infty$ for $j = 1, \ldots, n$. Therefore, (see Chung, 1967)

$$\lim_{n \to +\infty} \frac{E\{\max_{1 \leq k \leq n} \sum_{i=\beta_{k-1}+1}^{\beta_k} |f(Z_i)|\}}{n} = 0.$$

and

$$\lim_{n \to +\infty} \frac{E\{S_n\}}{n} = \lim_{n \to +\infty} \frac{E\{\sum_{j=1}^{l(n)} Y_j\}}{n}.$$

But

$$E\left\{\sum_{j=1}^{l(n)} Y_j\right\} = \sum_{m=1}^{n} \int_{l(n)=m} \sum_{\nu=1}^{m} Y_\nu \, dP,$$

$$= \sum_{\nu=1}^{n} \int_{l(n) \geq \nu} Y_\nu \, dP,$$

$$= \sum_{\nu=1}^{n} [E\{Y_\nu\} - \int_{l(n) < \nu} Y_\nu \, dP].$$

Now $\{z : l(n,z) < \nu\} = \{z : \beta_\nu > \beta_{l(n)}\} \in \mathcal{F}_{\beta_{\nu-1}}$. On the other hand, $\{z \mid Y_\nu = c\} \in \mathcal{F}'_{\beta_{\nu-1}}$. The independence of Y_ν and $1_{\{l(n) < \nu\}}$ is guaranteed by Theorem 11, and further simplification yields

$$E\left\{\sum_{j=1}^{l(n)} Y_j\right\} = \sum_{\nu=1}^{n} [E\{Y_\nu\} - P(l(n) < \nu) E\{Y_\nu\}]$$

$$= E\{Y\} \sum_{\nu=1}^{n} P(l(n) > \nu)$$

$$= E\{Y\} E\{l(n)\}$$

Then

$$\lim_{n \to +\infty} \frac{E\left\{\sum_{j=1}^{l(n)} Y_j\right\}}{n} = \lim_{n \to +\infty} E\{Y\} E\{l(n)/n\}$$

The process $l(n)$ is an example of a discrete time counting process, so the Elementary Renewal Theorem guarantees that

$$\lim_{n\to+\infty} \frac{l(n)}{n} = \frac{1}{\mathrm{E}\{\alpha\}}.$$

Thus

$$\begin{aligned}
\lim_{n\to+\infty} \frac{\mathrm{E}\{S_n\}}{n} &= \lim_{n\to+\infty} \frac{\mathrm{E}\left\{\sum_{j=1}^{n} Y_j\right\}}{n} \\
&= \lim_{n\to+\infty} \mathrm{E}\{Y\}\mathrm{E}\{l(n)/n\} \\
&= \frac{\mathrm{E}\{Y\}}{\mathrm{E}\{\alpha\}} \\
&= \mathrm{E}\{f(Z')\} \quad \blacksquare
\end{aligned}$$

This sequence of theorems is an adaptation of random walk and countable Markov chain results (see Chung, 1967 and 1974). The close relationship between the continuous and discrete processes makes the final theorem of the section straightforward.

(16) THEOREM. If the process $\{Z_t, t \geq 0\}$ is a GSMO with a single set, it is a regenerative process.

Proof. The natural renewal process is the sequence of moments the GSMO leaves the single set. The random variables

$$\gamma(k) = \sum_{i=\beta_{k-1}+1}^{\beta_k} \min_{C_{ij}k_{X_j e_i}>0} \frac{C_{ij}}{k_{X_j e_i}}$$

represent the length of the cycles in continuous time. They are independent and identically distributed, implying that $\{\gamma_n, n \geq 1\}$ is a renewal process. The Strong Markov property (see Çinlar, p. 239) applies to $\{Z(t), t \geq 0\}$ and guarantees the second regenerative requirement is satisfied (see Çinlar 8.1.14 p. 239). $\square$

The distribution of each γ_i is nonarithmetic, since the distribution of each is, and $P(\beta_k - \beta_{k-1} < +\infty) = 1$, so the existing regenerative techniques are appropriate for GSMOs with single sets.

3.2. EXAMPLES

In this section, use of the regenerative technique is illustrated in the simulation of two GSMOPs with single sets. Both examples are queueing networks which have arisen in the computer modeling literature. The first is the single server cyclic queue introduced in Section 2.2. The second is a data management model with resource contention. The assumptions made about each model are those which ease the calculation of exact results in order that they may be compared with the simulation results. The most important of these is the assumption of servers whose service times have gamma distributions. The remainder of the section is devoted to a more complete presentation of the models and a report of the simulation results.

(1) Example. Single Server Cyclic Queue

The single server cyclic queue introduced in Example 2.2.16 has been used as part of a control variables method for the analysis of a demand paging computer system. The details of this type of computer system and the reasons for choosing a cyclic queue for the control process are not the principle interest here; an explanation of these matters can be found in Gaver and Shedler (1971). The effectiveness of regenerative methods in simulating queueing networks is our main concern.

Recall that the state space for the GSMOP can be adequately described by an ordered pair—one coordinate representing the number of jobs waiting or being served by server A and the other representing the same quantity for server B (see Figure 2.2.1). The two events are service completions and the clocks governing service times are gamma (2,1) for server A and gamma (3,1) for server B. The queueing discipline is first in, first out at both stations and the initial distribution is given by $\nu(A) = \mathrm{P}\big(((0,k),0,c),A\big)$, for any positive c. Each Z-state is therefore a quadruple—(jobs at A, jobs at B, service time

remaining at A, service time remaining at B). All speeds are unity.

We will consider five functions $f{:}R^4 \mapsto R$ for this model. They are:

$$f_1(x) = 1_{\{0\}}(x_1)$$
$$f_2(x) = 1_{\{1\}}(x_1)$$
$$f_3(x) = 1_{\{2\}}(x_1)$$
$$f_4(x) = 1_{\{3\}}(x_1)$$
$$f_5(x) = 3 \cdot 1_{\{3\}}(x_1) + 2 \cdot 1_{\{2\}}(x_1) + 1_{\{3\}}(x_1).$$

The first four functions are used to estimate the stationary distribution of the number of jobs in service or waiting at station A. The fifth function is used to estimated the expected number of jobs at station A when the station is in equilibrium.

The simulation results for this model are displayed in 3 tables. Table 1 contains the results of runs using f_1 and f_2, Table 2 uses f_3 and f_4, and Table 3 uses f_5. In each table we list the point estimate and half length of a 90 percent confidence interval based on various numbers of cycles. For all our runs the cycles were based on returns to the state (0,k). Each table also includes an estimate $(\hat{\sigma}/\overline{a})$ of the variance term in the central limit theorem.

For these runs, all of the confidence intervals based on 1000 cycles covered the true value of the parameter estimated. Also, all of the variance estimates were within 5 percent of their true value.

TABLE 1

SIMULATION OF SINGLE SERVER CYCLIC QUEUE

3 jobs - $\Gamma(2,1)$, $\Gamma(3,1)$ servers

90 percent confidence intervals

$f_1(x) = 1_{\{0\}}(x_1)$			
Cycles	Point Estimate	Half Length	$\hat{\sigma}/\bar{\alpha}$
100	.0436	.0107	.1218
300	.0484	.0070	.1486
500	.0464	.0052	.1438
1000	.0428	.0035	.1339
THEORY	.0426		.1404

$f_2(x) = 1_{\{1\}}(x_1)$			
Cycles	Point Estimate	Half Length	$\hat{\sigma}/\bar{\alpha}$
100	.1752	.0226	.5442
300	.1862	.0153	.6010
500	.1846	.0106	.5925
1000	.1796	.0072	.5465
THEORY	.1767		.5286

TABLE 2

SIMULATION OF SINGLE SERVER CYCLIC QUEUE

3 jobs - $\Gamma(2,1)$, $\Gamma(3,1)$ servers

90 percent confidence intervals

$f_3(x) = 1_{\{2\}}(x_1)$			
Cycles	Point Estimate	Half Length	$\hat{\sigma}/\bar{a}$
100	.4226	.0193	.4004
300	.4094	.0108	.3511
500	.4138	.0080	.3324
1000	.4164	.0056	.3325
THEORY	.4189		.3464

$f_4(x) = 1_{\{3\}}(x_1)$			
Cycles	Point Estimate	Half Length	$\hat{\sigma}/\bar{a}$
100	.3409	.0213	.5107
300	.3530	.0147	.6441
500	.3582	.0101	.6159
1000	.3584	.0082	.7396
THEORY	.3617		.7579

TABLE 3

SIMULATION OF SINGLE SERVER CYCLIC QUEUE

3 jobs - $\Gamma(2,1)$, $\Gamma(3,1)$ servers

90 percent confidence intervals

$f_3(x) = 3 \cdot 1_{\{3\}}(x_1) + 2 \cdot 1_{\{2\}}(x_1) + 1_{\{1\}}(x_1)$			
Cycles	Point Estimate	Half Length	$\hat{\sigma}/\bar{\alpha}$
100	2.176	.0482	3.175
300	2.111	.0328	3.942
500	2.109	.0262	3.926
1000	2.105	.0176	3.548
THEORY	2.099		3.428

(2) **Example.** Data Base Management System

The second example used to illustrate regenerative techniques in queueing networks is a data management facility known as Data Language/I (DL/I). This model was developed by Lavenberg and Shedler (1976) and is depicted in Figure 4. A fixed number of jobs, each representing a data base call, circulate in the system. Note there can be more than one queue for a type of service and at the completion of service more than one route may be possible. The choice of route is determined by Bernoulli random variables ψ_1 and ψ_2.

The diagram differs from one of a conventional queueing network in that the circles represent services rather that servers. The α services $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ are rendered by a single server representing the central processing unit, while the β service is performed by a single server representing an I/O device. It is assumed that no two α services can be performed concurrently but that a β service may be performed at the same time as any α service. The services $\alpha_0, \alpha_1, \alpha_2$ and β are noninterruptable, but the α_3 service is suspended at the completion of a β service; this interruption is of the preemptive resume type.

A processor scheduling decision must be made at the completion of any α service or any β service when either no α service is being rendered or α_3 service is being performed. It is assumed that the customer whose service has been complete enters the next queue immediately. The next processor service is the service having the highest priority, where priority is determined by an ordering of the queues $q_0, q_{1,1}, q_{2,1}, q_{1,2}, q_{2,2}$ and q_3. The processor scheduling algorithm employed in this model is $q_{1,2}, q_{2,1}, q_{1,1}, q_{2,2}, q_0$ and q_3 from highest to lowest priority.

One way to model this system as a GSMOP is to define the state space as a 9-tuple. A coordinate is needed to record the number of jobs waiting for service at each of the β and six α waiting rooms. Also a coordinate is needed to record which waiting room the α server is servicing (numbered from 1 to 6, in

descending priority), and an indicator is required to help the system remember if an interrupted α_3 service is awaiting resumption of service. Three clocks are required. One to record the amount of time remaining in the β service; one for the time remaining in a α_0, α_1 or α_2 service and one for the remaining α_3 service. The α_3 service must be separated from the other α services because of the preemptive resume interruption at completion of β service. After the interruption, a new service receives the α server's attention and the speed associated with the α_3 service is zero. It remains zero until all the jobs are waiting for α_3 service. The data reported in Tables 5, 6 and 7 is for the system with two jobs, gamma (2,1) distributions governing the α_0 and α_2 services and exponential (1) distributions governing the remaining services. The binary random variable ψ_1 is 1 with probability .1, while ψ_2 is 1 with probability .2. Suppose the chain is distributed initially as if a service completion has just occurred when all the jobs in the system were waiting for the α_3 service. That is, $\nu(A) = P\big(\big((0,0,0,0,0,0,2,6,0),0,0,c\big),A\big)$. Cycles for this example will be based on returns to the state where all jobs wait for α_3 service.

The functions $f: R^9 \mapsto R$ we will examine for this system are:

$$f_1(x) = 1_{\{0\}}(x_8)$$
$$f_2(x) = 1_{\{2\}}(x_8)$$
$$f_3(x) = 1_{\{5\}}(x_8)$$
$$f_4(x) = 1_{\{1,2\}}(x_1)$$
$$f_5(x) = 2 \cdot 1_{\{2\}}(x_7) + 1_{\{1\}}(x_7).$$

The first three functions are used to determine the probability that when the system is in equilibrium the α server services a job in waiting room 0, 2 and 5 respectively. The fourth function will be used to determine the probability the β server is busy and the fifth to determine the expected length of queue in waiting room q_3.

The simulation results for f_1 and f_2 are found in Table 5, the results for

f_3 and f_4 are in Table 6 and the results for f_5 are in Table 7. The format of the tables is the same as for the Tables for Example 1.

For each function the confidence intervals (90 percent) based on 500 cycles covered the parameter estimated. In each case except one (f_5) the error in the variance estimate was less than 5 percent of the true variance.

This example, as well as Example 1, tends to support the use of regenerative techniques with models of GSMOPs which have a single set.

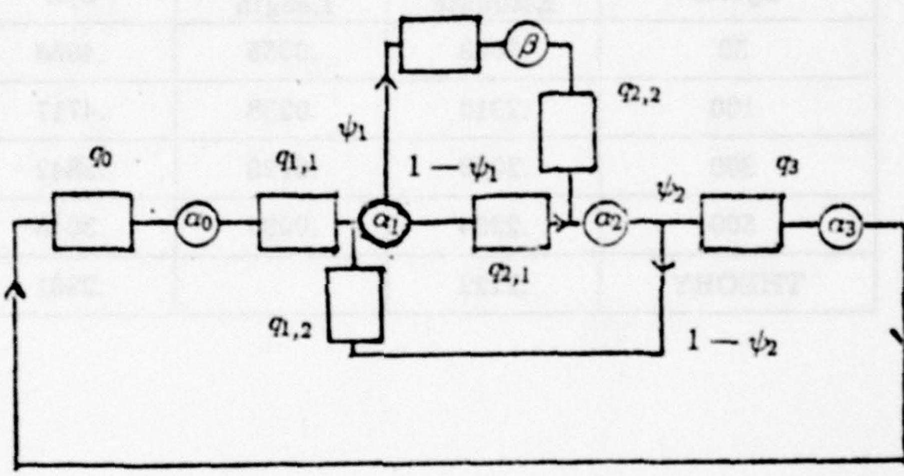

Figure 1. Data Management Model

TABLE 5

SIMULATION OF DATA BASE MANAGEMENT SYSTEM

2 jobs - Γ (2,1) service at α_0, α_2

90 percent confidence intervals

$f_1(x) = 1_{\{0\}}(x_8)$			
Cycles	Point Estimate	Half Length	$\hat{\sigma}/\overline{\alpha}$
50	.2003	.0355	.4088
100	.2310	.0238	.4717
300	.2253	.0129	.3842
500	.2254	.0095	.3048
THEORY	.2222		.2961

$f_2(x) = 1_{\{2\}}(x_8)$			
Cycles	Point Estimate	Half Length	$\hat{\sigma}/\overline{\alpha}$
50	.5242	.0294	.3761
100	.5089	.0227	.4228
300	.5137	.0129	.4298
500	.5062	.0104	.4533
THEORY	.4999		.4505

TABLE 6

SIMULATION OF DATA BASE MANAGEMENT SYSTEM

2 jobs - Γ (2,1) service at α_0, α_2

90 percent confidence intervals

$f_3(x) = 1_{\{5\}}(x_8)$			
Cycles	Point Estimate	Half Length	$\hat{\sigma}/\overline{a}$
50	.1094	.0243	.2563
100	.10484	.0164	.2208
300	.1078	.0092	.2163
500	.1117	.0074	.2281
THEORY	.1111		.2222

$f_4(x) = 1_{\{0\}}(x_1)$			
Cycles	Point Estimate	Half Length	$\hat{\sigma}/\overline{a}$
50	.0111	.0063	.0133
100	.0232	.0065	.0362
300	.0292	.0045	.0582
500	.0278	.0036	.0562
THEORY	.0277		.0538

TABLE 7

SIMULATION OF DATA BASE MANAGEMENT SYSTEM

2 jobs - Γ (2,1) service at α_0, α_2

90 percent confidence intervals

$f_5(x) = 2 \cdot 1_{\{2\}}(x_7) + 1_{\{1\}}(x_7)$				
Cycles	Point Estimate	Half Length	$\hat{\sigma}/\bar{\alpha}$
50	.9092	.0952	3.051
100	.8211	.0831	5.831
300	.8061	.0439	5.493
500	.8360	.0328	4.783
THEORY	.8526		4.055

CHAPTER IV

GSMOS WITHOUT SINGLE SETS

4.1. INTRODUCTION

UNFORTUNATELY, THE REQUIREMENT that a GSMP have a single state means that the results of Chapter III cannot be used to model many stochastic systems. As a system becomes more complex, it becomes increasingly unlikely that a state is associated with a only one event. The purpose of this chapter is to examine the possibilities for simulating these processes. Since GSMOPs are Markov chains the results presented in Section 2.1 are the principal tools available. A recurrence condition is required for the limit results reported in Section 2.1, so Section 4.2 explores the class of GSMOPs which satisfy the various conditions. All that is required for a GSMOP to be recurrent is that the densities governing the clocks of the process must have support on the positive half line. This allows the application of the regeneration and skeleton lemmas (2.1.9 and 2.1.11, respectively) of the Athreya and Ney presentation and makes possible the construction of a regenerative process closely connected to the original GSMOP. The central limit theorem associated with this regenerative process can be used to obtain the estimates required for the original process. This regenerative process and its central limit theorem are

developed in Section 4.4. The transition function for this regenerative process cannot be explicitly determined usually, so it is not satisfactory for simulation applications. To remedy this difficulty a way of developing 'quasi-cycles' in the original process is proposed. These quasi-cycles are based on a sequence of Bernoulli random variables which are *independent* of the original process and at the same time closely linked to the regenerative process. The new process based on quasi-cycles and its central limit theorem are also introduced in Section 4.4.

A nice feature of the quasi-cycle process is that the variance constant in its central limit theorem can be estimated. This problem is considered in Section 4.5 and a sample simulation is presented in Section 4.6.

An interesting ergodic theorem for recurrent GSMOs is developed in Section 4.3. It is a strong law with the limiting quantity in terms of the invariant measure for the associated GSMOP.

Throughout this chapter all speeds will be assumed to be unity. Most of the results are probably true without this assumption, but the introduction of speeds, particularly zero speeds, tends to complicate the situation.

4.2. RECURRENCE CONDITIONS

SINCE MOST COMPLEX systems do not have single states, identifying those GSMOPs which satisfy a *recurrence condition* becomes an important problem. Example 3.1.2 demonstrates that Doeblin's condition is too strong for many simple systems. The difficulty illustrated in this example arises in all GSMOPs which satisfy $|E(s)| > 2$ for all $s \in S$, as well as many systems where $|E(s)| \leq 2$ for some s. The problem is that some Z-states have clocks that, with a high probability, remain active for a large number of steps. This must be prevented if Doeblin's condition is to be satisfied because it allows sets with small φ-measure to be visited with very high probability by a particular Z-state. One way to do this is to suppose that $a_{s,e,e'} \leq T < +\infty$ for all choices of s e and e'. This guarantees that, for any $\delta > 0$, there is a finite integer m such that for all $(x, c) \in \Omega'$,

$$P_{(x,c)}(E(x) \subseteq \bigcup_{j=0}^{m} e_j^*) > 1 - \delta.$$

In other words, all the clocks initially active (those positive components of c) expire in the first m steps of the chain with a high probability. This is enough to guarantee Doeblin's condition.

(1) THEOREM. Let Z be a GSMOP satisfying $a_{s,e,e'} \leq T < +\infty$ for all $s \in S, e \in E(s), e' \in E$. Then Z satisfies Doeblin's condition.

Proof. Observe that from any state (x,c) in Ω', there is an i in the set $(1, 2, \ldots, |E(x)| + 1)$ where $e_i^* \not\in E(x)$. That is, in the first $|E(x)| + 1$ steps, at least one event that occurs is *not* originally scheduled in x. This implies that the length of the first $k \cdot \max_{x \in S} E(x) + k$ steps of a GSMOP can be bounded below by the sum of the k random variables. Each of these random variables is chosen from the family $\{ Y_{s,e,e'} : s \in S, e \in E(s), e' \in E \}$. (Recall that when event e' occurs and the new state is s, the clock associated with

event $e \in E(s)$ is set according to the distribution $F(\cdot, s, e, 0, e')$. $Y_{s,e,e'}$ is a random variable which has this distribution). In addition, for any positive ϵ, there is a positive integer j such that the sum of j independent realizations of $Y_{s,e,e'}$ exceeds $\max_{s,e,e'} a_{s,e,e'}$ with a probability of at least $1 - \epsilon$. These two facts imply that

$$P_{(x,c)}(Q_{j \cdot \max|E(x)|+j} > \max_{s,e,e'} a_{s,e,e'}) > 1 - \epsilon \tag{2}$$

for all $(x, c) \in \Omega'$. Let $q = j \cdot \max_{x \in S}|E(x)| + j$. (Note that q and j have ϵ as an argument but we will suppress it.) The set $(Q_q > \max_{s,e,e'} a_{s,e,e'})$ is a subset of $((x, c){:}E(x) \subseteq \bigcup_{j=0}^{q} e_j^{\bullet})$ so the latter set has a probability exceeding $1 - \epsilon$ as well.

For $B \in \mathfrak{F}'$, define $B_i = ((x, c){:}x = s_i)$ for $i = 1, 2, \ldots, |S|$, and note that $B = \bigcup_{i=1}^{|S|} B_i$ and the B_i's are disjoint. It must be true that

$$\begin{aligned} P^q((x,c),B) &= P^q((x,c),B \mid E(x) \subseteq \textstyle\bigcup_{j=0}^{q} e_j^{\bullet})\, P^q_{(x,c)}(E(x) \subseteq \textstyle\bigcup_{j=0}^{q} e_j^{\bullet}) \\ &\quad + P^q((x,c),B \mid E(x) \not\subseteq \textstyle\bigcup_{j=0}^{q} e_j^{\bullet})\, P^q_{(x,c)}(E(x) \not\subseteq \textstyle\bigcup_{j=0}^{q} e_j^{\bullet}) \\ &\leq \sum_{i=1}^{|S|} P^q((x,c),B_i \mid E(x) \subseteq \textstyle\bigcup_{j=0}^{q} e_j^{\bullet}) + \epsilon. \end{aligned} \tag{3}$$

A reference measure must now be constructed to take advantage of the decomposition demonstrated in (3). Let m_i be Lebesgue measure on R^i. With a slight abuse of notation, let us write $m_i(B_i)$ for the Lebesgue measure of the projection of the positive components of B_i's clocks on $R^{|E(s_i)|}$. Since we are assuming that $a_{s,e,e'} \leq T$, we have $m_i(C(s_i)) \leq T^{|E(s_i)|}$.

The condition on each summand in (4) is an important one since it means that all the clocks active in s_i were activated at some step in the path from x to s_i. This prevents the situation illustrated by Example 3.1.2 when Doeblin's condition was not satisfied. Therefore we must have that

$$P^q((x,c),B_i \mid E(x) \subseteq \textstyle\bigcup_{j=0}^{q} e_j^{\bullet}) \leq K^{|E(s)|} m_i(B_i). \tag{4}$$

Define

$$M(B) = \frac{1}{|S|} \sum_{i=1}^{|S|} \frac{m_i(B_i)}{m_i(C(s_i))}.$$

Recall that $C(s_i)$ is the collection of all possible clock readings when the process is in state s_i. Note that M is a probability measure on Ω'. It is nonnegative and the empty set has measure zero. For pairwise disjoint $D_i \in \Omega'$, $i = 1, 2, \ldots,$ define $D_{i,j} = ((x, c) \in D_i : x = s_j)$. Then

$$M(\bigcup_{i=1}^{\infty} D_i) = \frac{1}{|S|} \sum_{j=1}^{|S|} \frac{m_j(\bigcup_{i=1}^{\infty} D_{ij})}{m_j(C(s_j))},$$

$$= \frac{1}{|S|} \sum_{j=1}^{|S|} \sum_{i=1}^{\infty} \frac{m_j(D_{ij})}{m_j(C(s_j))};$$

$$= \sum_{i=1}^{\infty} \frac{1}{|S|} \sum_{j=1}^{|S|} \frac{m_j(D_{ij})}{m_j(C(s_j))};$$

$$= \sum_{i=1}^{\infty} M(D_i)$$

Furthermore, it must be that

$$m_i(B_i) \leq m_i(C(s_i)) \cdot |S| \cdot M(B). \tag{5}$$

From (3), (4) and (5) we have

$$
\begin{aligned}
P^q((x,c),B) &\leq \sum_{i=1}^{|S|} P^q((x,c),B_i \mid E(x) \subseteq \textstyle\bigcup_{j=0}^{q} e_j^{*}) + \epsilon \\
&\leq \sum_{i=1}^{|S|} K^{|E(s_i)|} m_i(B_i) + \epsilon \\
&\leq \sum_{i=1}^{|S|} K^{|E(s_i)|} m_i(C(s_i))\,|S|\,M(B) + \epsilon \\
&\leq K^{|E(s_i)|} \max_{1 \leq i \leq |S|} m_i(C(s_i))\,M(B)\,|S|^2 + \epsilon \\
&\leq (KT)^{\max|E(s_i)|}\,|S|^2\,M(B) + \epsilon.
\end{aligned}
$$

There is a $\delta > 0$ such that $2\delta < 1 - \delta/(KT)^{\max|E(s_i)|}|S|^2 M(B)$. Pick $\epsilon = \delta$, then

$$M(B) < \delta/(KT)^{\max|E(s_i)|}|S|^2 M(B)$$

implies

$$P^q((x,c),B) < 2\delta < 1 - \delta/(KT)^{\max|E(s_i)|}|S|^2 M(B).$$

The chain therefore satisfies Doob's Condition. ∎

For those GSMOPs whose clocks have positive density on $[0,\infty)$, there are many Doeblin recurrent chains which are closely related. Define a new chain V in the following manner. Let the state space S and the family of transition functions $\{\, p(s,e,e') \,\}$ be the same in the two chains. Change only the densities of the clocks. Choose and $U \in R^+$ and define the new densities $g_{s,e,e'}$ for V by

$$g_{s,e,e',U}(x) = \begin{cases} f_{s,e,e'}(x)/F(U,s,e,0,e'), & \text{if } 0 < x \leq U; \\ 0, & \text{if } U < x. \end{cases} \tag{7}$$

Then V satisfies Doeblin's condition and is $(\Omega'_V, \xi, \lambda, n)$-recurrent, for some ξ, λ, and n. We shall call V the truncated Markov chain of Z. Using this truncated chain, we can show that Z is a Harris recurrent chain.

(8) **THEOREM.** Let Z be a GSMOP with $a_{s,e,e'} = +\infty$ for all $s \in S$, $e \in E(s)$ and $e' \in E$. Then Z is (A, φ, λ, n)-recurrent for some choice of A, φ, λ and n.

Proof. Choose a $U \in R^+$ and form the truncated chain V using the densities in (6). For $A \in \mathcal{F}'_V$, $z_0 \in \Omega'_V$, we claim that, for $n = 1, 2, \ldots$,

$$P^n_Z(z_0, A \mid z_i \in \Omega'_V, i = 1, 2, \ldots n) = P^n_V(z_0, A).$$

Since both P^n_Z and P^n_V are finite, we need only consider sets A of the form

$$B = \{s'\} \times \{\times_{j=1}^{|E|}[0, \ell_j)\}$$

where $s' \in S$ and $\ell_j = 0$ if $e_j \notin E(s)$, $U \geq \ell_j > 0$ otherwise. We will prove the claim by induction. Suppose n=1.

$$\begin{aligned} P_V((s,c), A) &= p(s, e^*(s,c), s') \prod_{j=1}^{|E(s')|} F_V(\ell_j, s', e_j, t_j, e') \\ &= p(s, e^*(s,c), s') \prod_{\{j:t_j>0\}} F_V(\ell_j, s', e_j, t_j, e') \prod_{\{j:t_j=0\}} F_V(\ell_j, s', e_j, t_j, e') \end{aligned}$$

The functions in the first product are all indicator functions and once (s,c) and B are given they are fixed. The functions in the second product are those clocks which are turned on when the process reaches s' after leaving s. To prove our claim, we need only worry about the distributions in the second product, all other terms in the expression are fixed. We can rewrite each term in the second product

$$\begin{aligned} F_V(\ell_j, s', e_j, 0, e') &= P_V(Y_{s,e,e'} \leq \ell_j) \\ &= P_Z(Y_{s,e,e'} \leq \ell_j)/P_Z(Y_{s,e,e'} \leq U) \\ &= P_Z(Y_{s,e,e'} \leq \ell_j \mid Y_{s,e,e'} \leq U) \\ &= P_Z(Y_{s,e,e'} \leq \ell_j \mid Y_{s,e,e'} \leq U \text{ for all } e \in E(s)). \end{aligned}$$

The last equality is true since the Y's are all independent. Therefore the assertion is true for n=1.

Now suppose it is true for $n = 1, \ldots, m-1$. We must show it is true for $n = m$. We can write

$$\mathrm{P}_Z^m(z_0, B \mid z_i \in \Omega'_V, i = 1, \ldots, m)$$

$$= \int_{\Omega'_V} \mathrm{P}_Z^{m-1}(z_1, B \mid z_i \in \Omega'_V, i = 1, \ldots, m)\mathrm{P}(z_0, dz_1 \mid z_i \in \Omega'_V, i = 1, \ldots, m)$$

$$= \int_{\Omega'_V} \mathrm{P}_Z^{m-1}(z_1, B \mid z_i \in \Omega'_V, i = 1, 2, \ldots, m-1)\,\mathrm{P}(z_0, dz_1 \mid z_1 \in \Omega'_V)$$

$$= \int_{\Omega'_V} \mathrm{P}_V^{m-1}(z_1, B)\mathrm{P}_V(z_0, dz_1).$$

We conclude the proof be choosing n so that V is $(\Omega'_V, \varphi, \lambda, n)$-recurrent for some φ and λ. Then, for $z \in \Omega'_V$, $A \subseteq \Omega'_V$,

$$\begin{aligned} \mathrm{P}_Z^n(z, B) &\geq \mathrm{P}_Z^n(z, B \mid z_i \in \Omega'_V, i = 1, \ldots, n) \\ &= \mathrm{P}_V^n(z, B) \\ &\geq \lambda\varphi(A). \end{aligned}$$

Therefore Z is $(\Omega'_V, \varphi, \lambda, n)$-recurrent. ∎

4.3. ERGODIC THEORY FOR GSMOS

THE DEVELOPMENT OF an ergodic theorem for GSMOs associated with recurrent GSMOPs is the topic of this section. The main results are Theorems 1 and 7. The first is a limit theorem for the counting process generated by a GSMO. The second is a continuous time strong law that relates the limiting behavior of a GSMO to the invariant measure of its associated GSMOP. Throughout this section, it will be assumed that the GSMOP Z is recurrent and has π as its invariant measure. Define a function $f:\Omega' \to R$ by

$$f(x, c) = \min_{c_i > 0} c_i \cdot g(x)$$

where g is a measurable function from S to R. Let

$$\mu = \mathrm{E}_\pi\{ \min_{C_{0i} > 0} C_{0i} \}$$

and

$$\kappa = \mathrm{E}_\pi\{ f(Z_0) \}.$$

(1) THEOREM. Let Z be a GSMO with counting process N(t). Then

$$\lim_{t \to +\infty} \frac{N(t)}{t} = \frac{1}{\mu} \qquad \mathrm{P}_\nu - a.e.$$

for any probability measure ν.

Proof. By the definitions of N(t) and Q(n) (see Section 2.2)

$$Q(N(t)) < t < Q(N(t) + 1)$$

or

$$\frac{Q(N(t))}{N(t)} < \frac{t}{N(t)} < \frac{Q(N(t)+1)}{N(t)} \tag{2}$$

when $N(t) \geq 1$. Since $Q(n) = \sum_{i=0}^{n-1} \min_{C_{ij}>0} C_{ij}$, the Strong Law of Large Numbers for Markov chains (Theorem 2.1.7) implies

$$\lim_{n \to +\infty} \frac{Q(n)}{n} = \mathrm{E}_\pi\{ \min_{C_{0i}>0} C_{0i} \}$$
$$= \mu \quad \mathrm{P}_\nu - a.e.$$

for any probability measure ν. The Strong Law's Hypotheses are satisfied since

$$\kappa < \max_{s,e,e'} \mathrm{E}\{ Y(s,e,e') \} \max_{s \in S} g(s) < +\infty.$$

Since $\lim_{t \to +\infty} N(t) = +\infty$,

$$\lim_{t \to +\infty} \frac{Q(N(t))}{N(t)} = \mu \quad \mathrm{P}_\nu - a.e. \tag{3}$$

and

$$\lim_{t \to +\infty} \frac{Q(N(t)+1)}{N(t)} = \lim_{t \to +\infty} \frac{Q(N(t)+1)}{N(t)+1} \cdot \frac{N(t)+1}{N(t)} = \mu \quad \mathrm{P}_\nu - a.e. \tag{4}$$

for any probability measure ν. The result is then immediate from (2), (3) and (4). ∎

The following two theorems lay groundwork for the ergodic theorem.

(5) THEOREM. Let Z be a GSMOP with invariant measure π. Then

$$\lim_{n \to +\infty} \frac{|f(Z_n)|}{n} = 0 \quad \mathrm{P}_\nu - a.e.$$

for any probability measure ν.

Proof. The Strong Law implies

$$\lim_{n \to +\infty} \frac{1}{n+1} \sum_{i=0}^{n} |f(Z_i)| = \lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} |f(Z_i)|$$
$$= \mathrm{E}_\pi\{ |f(Z_i)| \} \quad \mathrm{P}_\nu - a.e.$$

Then

$$\lim_{n\to+\infty}\frac{|f(Z_n)|}{n}=\lim_{n\to+\infty}\frac{1}{n}\Big(\sum_{i=0}^{n}|f(Z_i)|-\sum_{i=0}^{n-1}|f(Z_i)|\Big),$$

$$=\lim_{n\to+\infty}\frac{n+1}{n}\lim_{n\to+\infty}\frac{\sum_{i=0}^{n}|f(Z_i)|}{n+1}-\lim_{n\to+\infty}\frac{\sum_{i=0}^{n-1}|f(Z_i)|}{n}$$

$$=0 \quad \mathrm{P}_\nu - a.e.$$

for any probability measure ν. ∎

(6) **THEOREM.** Let $\mathfrak{X}=(\mathfrak{X},\mathfrak{C})$ be a GSMO whose associated GSMOP is recurrent. Then

$$\lim_{t\to+\infty}\frac{\int_0^{Q(N(t)+1)}g(\mathfrak{X}_t)\,dt-\int_0^t g(\mathfrak{X}_t)\,dt}{t}=0 \qquad \mathrm{P}_\nu - a.e.$$

for any probability measure ν.

Proof.

$$\lim_{t\to+\infty}\frac{\left|\int_0^{Q(N(t)+1)}g(\mathfrak{X}_t)\,dt-\int_0^t g(\mathfrak{X}_t)\,dt\right|}{t},$$

$$=\lim_{t\to+\infty}\frac{\left|\int_0^{Q(N(t)+1)}g(\mathfrak{X}_t)\,dt-\int_0^{Q(N(t))}g(\mathfrak{X}_t)\,dt-\int_{Q(N(t))}^t g(\mathfrak{X}_t)\,dt\right|}{t},$$

$$\le\lim_{t\to+\infty}\frac{\left|\int_0^{Q(N(t)+1)}g(\mathfrak{X}_t)\,dt-\int_0^{Q(N(t))}g(\mathfrak{X}_t)\,dt\right|}{t}$$

$$+\lim_{t\to+\infty}\frac{\left|\int_{Q(N(t))}^t g(\mathfrak{X}_t)\,dt\right|}{t},$$

$$\le\lim_{t\to+\infty}\frac{2\int_{Q(N(t))}^{Q(N(t)+1)}|g(\mathfrak{X}_t)|\,dt}{N(t)}\frac{N(t)}{t},$$

$$=2\lim_{t\to+\infty}\frac{|f(Z_{N(t)})|}{N(t)}\lim_{t\to+\infty}\frac{N(t)}{t},$$

$$=0.$$

The last equality is true from Theorem 5 and the fact that $\lim_{t\to\infty}N(t)=+\infty$. ∎

The next result in this section is a strong law for $\mathfrak{Z}$. It is unusual in that the limiting result is in terms of a discrete process, Z.

(7) THEOREM. Let $\mathfrak{Z}$ be a GSMO whose associated GSMOP is recurrent. Then

$$\lim_{t\to+\infty} \frac{\int_0^t g(\mathfrak{Z}_t)\,dt}{t} = \frac{\kappa}{\mu} \quad \mathrm{P}_\nu - a.e.$$

for any probability measure ν.

Proof. It is clear that

$$\frac{\int_0^t g(\mathfrak{Z}_t)\,dt}{t} = \frac{\int_0^{Q(N(t)+1)} g(\mathfrak{Z}_t)\,dt - \int_0^{Q(N(t)+1)} g(\mathfrak{Z}_t)\,dt + \int_0^t g(\mathfrak{Z}_t)\,dt}{t},$$

$$= \frac{\sum_{n=0}^{N(t)} f(Z_n)}{N(t)+1}\,\frac{N(t)+1}{t} - \frac{\int_0^{Q(N(t)+1)} g(\mathfrak{Z}_t)\,dt - \int_0^t g(\mathfrak{Z}_t)\,dt}{t},$$

Since $N(t) \to +\infty$ as $t \to +\infty$, the Strong Law of Large Numbers (Lemma 2.1.7) and Theorems 5 and 6 imply that

$$\lim_{t\to+\infty} \frac{\int_0^t g(\mathfrak{Z}_t)\,dt}{t} = \frac{\kappa}{\mu} \quad \mathrm{P}_\nu - a.e.$$

for any probability measure ν. ∎

4.4. SUPPLEMENTARY VARIABLES AND GSMOS

AT THIS POINT we are ready to tackle the problem of developing a central limit theorem for GSMOs which do not have a single set. We will attempt to exploit the recurrence properties of the associated GSMOP to construct a regenerative process that is very similar to the GSMO being studied. The important results for this approach are given in Lemmas 2.1.9 and 2.1.11. Using the techniques in the proof of Lemma 2.1.9, it is easy to find a regenerative process that is closely related to any GSMO when its associated GSMOP is aperiodic and (A, φ, λ, k)-recurrent. Unfortunately, this regenerative process is not well suited for simulation applications because part of its transition function usually cannot be determined. To avoid this difficulty, we propose a new approach to simulating the original process by using an independent sequence of Bernoulli random variables to break a sample path into 'quasi-cycles.' A central limit theorem based on these quasi-cycles will be developed in this section and the estimation of the its variance constant will be discussed in Section 4.5.

Suppose the Markov chain Z on $(\Omega', \mathcal{F}')$ is aperiodic and (A, φ, λ, k)-recurrent with transition function P and starting measure φ. From Lemma 2.1.9 and Lemma 2.1.11 it is possible to decompose P^k in the manner demonstrated in Equation 2.1.8, that is,

$$P^k(x, A) = p\varphi(A \cap B) + (1-p)Q_k(x, B).$$

Now define a new process

$$T = \{ T_n = (V_n, \delta_n, l_n, G_n) : n = 0, 1, \dots \}$$

on $\Xi = \Omega' \times \{0, 1\} \times \{1, 2, \dots, k\} \times \Omega'$. Suppose that

$$P'(T_0 \in (B, \delta, l, C)) = 1_{\{1\}}(\delta) 1_{\{k\}}(l) \varphi(A \cap B \cap C)$$

and let the transition function P′ of T be defined by

$$P'((v,\delta,j,y),(B,\gamma,l,C)) = 1_{\{\delta\}}(\gamma)1_{\{j-1\}}(l)1_{\{C\}}(y)P(v,B)$$
$$\text{for } j = 2,3,\ldots,k;\ v,y \in E;\ \delta = 0 \text{ or } 1.$$

$$P'((v,\delta,1,y),(B,1,l,C)) = \begin{cases} p1_{\{k\}}(l)P(v,B), & y \notin A, \\ p1_{\{k\}}(l)\varphi(A \bigcap B \bigcap C), & y \in A; \end{cases} \tag{1}$$

$$P'((v,\delta,1,y),(B,0,l,C)) = \begin{cases} (1-p)1_{\{k\}}(l)P(v,B), & y \notin A, \\ (1-p)1_{\{k\}}(l)Q_k(y,B \bigcap C), & y \in A. \end{cases}$$

This transition function breaks T into groups of k steps. If, at the beginning of a 'cycle' (that is, when $l = k$), V_n is not in A, then each of the next k transitions is determined by the transition function P of the original process Z. If, when $l = k$, V_n is an element of A, the k-cycle has two pieces. The first $k-1$ transitions are determined by P again, but the last transition in the cycle is either φ or Q_k according to whether the random variable δ_{n+k} is 1 or 0. If Q_k is chosen the process must remember where it was k steps in the past (G_n). This is the most recent state of the skeleton process. Note that when $l = k$, the first and fourth coordinates of the T process are the same. This resets the fourth coordinate which indicates the last state V's skeleton process has visited.

In a manner analogous to 2.1.9(i) and (ii) it can be shown that

(2) $\{V_n, n = 0,1,\ldots\}$ is a Markov chain with transition probability function P and initial measure φ; and

(3) $\{\delta_n, n = 0,1,\ldots\}$ is a sequence of independent and identically distributed Bernoulli random variables with parameter p.

It is easy to see using the reasoning of Theorems 3.1.11 thru 3.1.15 that a sample path of T can be broken into independent and identically distributed pieces and that functions of these pieces are also i.i.d.. The stopping times

required for these results would be those coordinates when V is distributed according to φ.

If the Markov chain $Z = (X, C)$ is also a GSMOP on Ω', then $V_n = (U_n, K_n)$, where $U_n \in S$ and $K_n \in C(U_n)$ for all n. A continuous version of T, say

$$\mathcal{T} = (\mathcal{V}, \delta, \ell, \mathcal{G}) = ((\mathcal{U}, \mathcal{K}), \delta, \ell, \mathcal{G}) = (((\mathcal{U}_t, \mathcal{K}_t), \delta_t, \ell_t, \mathcal{G}_t), t \geq 0),$$

can be constructed in the same way as the GSMO $\mathcal{Z} = (\mathcal{X}, \mathcal{C})$ is constructed. Using the reasoning developed in Theorems 3.1.11 through 3.1.16, it can be shown that $\mathcal{T}$ is a regenerative process with the regenerative times being those epochs when $\mathcal{V}$ is distributed according to φ. This means that $\mathcal{V}$ has a limiting random variable, say $\mathcal{V}' = (\mathcal{U}', \mathcal{K}')$. Furthermore, since V_n has the same distribution as Z_n (both are Markov chains on Ω', with the same starting measure and transition function), it must be that $\mathcal{V}_t = (\mathcal{U}_t, \mathcal{K}_t)$ has the same distribution as $\mathcal{Z}_t$, and therefore $\mathcal{V}'$ has the same distribution as $\mathcal{Z}'$.

A standard central limit theorem can be written for $\mathcal{T}$. Suppose

$$\beta_0(T) = 0$$

$$\beta_j(T) = \min(kn > \beta_{j-1}(T) : \delta_{kn} = 1, V_{k(n-1)} \in A, n \geq 1)$$

$$Y_k = \sum_{n=\beta_{k-1}(T)}^{\beta_k(T)-1} \min_{K_{nj}>0} K_{nj} g(U_n) \tag{4}$$

$$\alpha_k = \sum_{n=\beta_{k-1}(T)}^{\beta_k(T)} \min_{K_{nj}>0} K_{nj}$$

for $k = 1, 2, \ldots$. The central limit theorem is then, if $\mathrm{E}\{|g(\mathcal{U}')|\} < +\infty$ and $\sigma^2 = \mathrm{Var}\,(Y_k - \mathrm{E}\{g(\mathcal{U}')\}\alpha_k)$,

$$\frac{\sum_{k=1}^{n}(Y_k - \mathrm{E}\{g(\mathcal{U}')\}\alpha_k)}{\sigma\sqrt{n}} \Rightarrow N(0, 1). \tag{5}$$

We can substitute $E\{g(\mathfrak{X}')\}$ for $E\{g(\mathfrak{U}')\}$ and we do so. The technique for estimating σ in this central limit theorem would be the standard regenerative estimate given in Section 2.3. Let $\hat{r}_T$, s_{11}^T, s_{22}^T, s_{12}^T and s_T be the estimates based on T for the obvious quantities.

Theoretically, our construction is now complete. We have constructed a regenerative Markov chain for an arbitrary aperiodic, (A, φ, λ, k)-recurrent GSMOP Z which can be used to estimate functions of the limiting random variable of Z's associated GSMP. Unfortunately, in order to simulate $\mathcal{T}$ one needs to know φ and Q_k, or at least how to sample from them. The problem can usually be constructed so that φ is known. The measure Q_k presents a problem, however. We only know its form through its relationship with φ and P^k. In practice calculating an explicit expression for P^k is usually hopeless and this makes the outlook dim for the generation of Q_k. For this reason it is difficult to use $\mathcal{T}$ as a means of estimating functions of $\mathfrak{X}'$.

There are certain features of the T process which suggest other means of estimating $\mathfrak{X}'$ through simulation. As we have noted in (2) and (3), V is a Markov chain with transition function P and δ is a sequence of independent and identically distributed Bernoulli random variables with parameter p. It is important to note, however, that V and δ are not independent sequences. It is the values of the δ sequences which determine the choice of Q_k or φ at those transitions when the cycles of T begin and end. Suppose that instead of using correlated sequences V and δ, two *independent* sequences are used.

Suppose that at multiples of k steps during the simulation of the original process Z, an independent Bernoulli random variable with parameter p is generated. Call the variable generated at the nk^{th} step χ_{nk}. Suppose that the elements of the $\chi = \{\chi_{nk}, n = 0, 1, \dots\}$ sequence are mutually independent and are used to define 'quasi-cycles'. We will base a central limit theorem for $\mathfrak{X}$ on these quasi-cycles. Toward this end, define the new stochastic process

$\mathcal{Q} = (\mathcal{Q}_n = (Z_n, \chi_n, j_n), n = 0, 1, \ldots)$ with initial measure

$$P_{\mathcal{Q}}(\mathcal{Q}_0 \in (A, \chi, j)) = P(Z_0 \in A) 1_{\{1\}}(\chi) 1_{\{k\}}(j)$$

and probability transition function given by

$$P_{\mathcal{Q}}((z, \chi, j), (A, \gamma, \ell)) = P(z, A) 1_{\{\ell\}} 1_{\{\chi\}}(\gamma) \text{ for } j = 2, \ldots, k$$

$$P_{\mathcal{Q}}((z, \delta, 1), (A, \gamma, \ell)) = [p\gamma + (1 - p)(1 - \gamma)] 1_{\{k\}}(\ell) P(z, A).$$

Also define the random variables

$$\beta_0(\mathcal{Q}) = 0,$$

$$\beta_i(\mathcal{Q}) = \min\{\, nk > \beta_{i-1}(\mathcal{Q}) \mid \chi_{nk} = 1, Z_{k(n-1)} \in A, n \geq 1 \,\},$$

$$L_k = \sum_{n = \beta_{k-1}(\mathcal{Q})}^{\beta_{k-1}(\mathcal{Q})} \min_{C_{nj} > 0} C_{nj} g(X_n),$$

$$\upsilon_k = \sum_{n = \beta_{k-1}(\mathcal{Q})}^{\beta_k(\mathcal{Q}) - 1} \min_{C_{nj} > 0} C_{nj}.$$

At this point, it is technically convenient for us to suppose that Z is Doeblin recurrent instead of Harris recurrent. To do this we will suppose that the densities are truncated at U in the manner we discussed in Section 4.4. This means that the length of the quasi-cycles are determined solely by the Bernoulli random variables since $A = \Omega'$.

(7) **THEOREM.** Let Z be an aperiodic, $(\Omega', \varphi, \lambda, k)$-recurrent GSMOP whose associated GSMO $\mathcal{Z} = (\mathcal{S}, C)$ has a weak limit $\mathcal{Z}' = (\mathcal{S}', C')$. If g is a measurable function from S to R and $E\{\,|g(\mathcal{S}')|\,\} < \infty$, then

$$\frac{\sum_{i=0}^{n-1} (L_i - E\{\, g(\mathcal{S}') \,\} \upsilon_i)}{\sigma \sqrt{n}} \Rightarrow N(0, 1)$$

where σ is defined in (4).

Proof. The proof will rely heavily on the central limit theorem for the T process, so we begin by studying that theorem. Define

$$R(k) = \max\{\, j \mid \beta_j(T) \le k \,\}.$$

Note that $R(\beta_k(T)) = k$ and that $R(n) = k$ if and only if $\beta_k(T) < n \le \beta_{k+1}(T)$. Now rewrite the central limit theorem for T as

$$\frac{\sum_{m=0}^{\beta_n(T)-1}(\min_{K_{mi}>0} K_{mi} f(U_m) - \mathrm{E}\{\, g(\mathfrak{S}')\,\} \min_{K_{mi}>0} K_{mi})}{\sigma\sqrt{n}} \Rightarrow N(0,1).$$

The random sum can be replaced by a fixed sum by noticing that

$$\frac{\sum_{m=0}^{\beta_{R(n)}(T)}(\min_{K_{mi}>0} K_{mi}\, g(U_m) - \mathrm{E}\{\, g(\mathfrak{S}')\,\} \min_{K_{mi}>0} K_{mi})}{\sigma\sqrt{R(n)}}$$

$$+ \frac{\sum_{m=\beta_{(R(n)+1)}(T)}^{n}(\min_{K_{mi}>0} K_{mi} g(U_m) - \mathrm{E}\{\, g(\mathfrak{S}')\,\} \min_{K_{mi}>0} K_{mi})}{\sigma\sqrt{R(n)}}$$

$$= \frac{\sum_{m=0}^{n}(\min_{K_{mi}>0} K_{mi}\, g(U_m) - \mathrm{E}\{\, g(\mathfrak{S}')\,\} \min_{K_{mi}>0} K_{mi})}{\sigma\sqrt{R(n)}} \tag{8}$$

and, for any $\epsilon > 0$,

$$\left| \frac{\sum_{m=\beta_{R(n)}}^{n}(\min_{K_{mi}>0} K_{mi} g(U_m) - \mathrm{E}\{\, g(\mathfrak{S}')\,\} \min_{K_{mi}>0} K_{mi})}{\sigma\sqrt{R(n)}} \right|$$

$$\le \frac{\sum_{m=\beta_{R(n)}(T)}^{n} \min_{K_{mi}>0} K_{mi} |g(U_m)| + \mathrm{E}\{\, |g(\mathfrak{S}')| \,\} \min_{K_{mi}>0} K_{mi}}{\sigma\sqrt{R(n)}}$$

$$\le \frac{\max_{s\in S} 2|f(s)|}{\sigma} \, \frac{\sum_{m=\beta_{R(n)}(T)}^{\beta_{(R(n)+1)}(T)} \min_{K_{mi}>0} K_{mi}}{\sqrt{R(n)}} \tag{9}$$

Since $(\beta_{(R(n)+1)}(T) - \beta_{(R(n)}(T)) < \infty$ and $Y_{\epsilon,e,e'} < +\infty$ with probability 1, it must be true that

$$\lim_{l \to +\infty} \mathrm{P}\left(\frac{\max f(s)}{\sigma} \; \frac{\sum_{m=\beta_{R(n)}(T)}^{\beta_{(R(n)+1)}(T)} \min_{K_{mi}>0} K_{mi}}{\sqrt{R(n)}} \leq \epsilon \right) = 1. \qquad (10)$$

Equations (8), (9) and (10) imply

$$\frac{\sum_{m=0}^{n}(\min_{K_{mi}>0} K_{mi} g(U_n) - \mathrm{E}\{g(\mathfrak{X}')\} \min_{K_{mi}>0} K_{mi})}{\sigma\sqrt{R(n)}} \Rightarrow N(0,1).$$

when used in conjuction with the so-called 'Converging Together' Lemma (see Chung (1974) p. 92 or Billingsley (1968) p. 25) and the fact that $R_j \to \infty$ as $j \to \infty$.

We have noted that V is a Markov chain with transition function P and initial measure ν. The GSMOP Z is also such a chain. This implies that $g(U_j)$ and $\min_{K_{mi}>0} K_{mi}$ can be replaced by $g(X_j)$ and $\min_{C_{mi}>0} C_{mi}$, respectively. Now define

$$W(k) = \max\{\, j : \beta_j(\Omega) \leq k \,\}.$$

Since X is $(\Omega', \varphi, \lambda, k)$-recurrent (here is the point where the assumption of Doeblin recurrence makes things a bit easier), the sequences δ and χ are both sequences of independent and identically distributed Bernoulli random variables with parameter p. This implies that

$$\frac{W(k)}{R(k)} \to 1 \quad \text{with probability 1}$$

and W(k) can replace R(k) in the central limit theorem. This yields

$$\frac{\sum_{m=0}^{n} \min_{C_{mi}>0} C_{mi} g(X_j) - \mathrm{E}\{g(\mathfrak{X}')\} \min_{C_{mi}>0} C_{mi}}{\sigma\sqrt{W(n)}} \Rightarrow N(0,1).$$

In the same way that the random sum for the V process was replaced by a nonrandom sum above, the nonrandom sum for X can be replaced by a random sum and the theorem is proven. ∎

The only obstacle remaining to our use of the independent sequences Z and χ is an explanation of how to estimate σ. This topic is important enough to be dealt with separately. We shall address it in Section 4.5.

4.5. VARIANCE ESTIMATION

IN MANY SIMULATION problems, particularly those concerned with determining confidence intervals, the estimation of a variance constant is one of the trickiest problems that must be resolved. Our situation is no exception. The variance constant in 4.4.5 is for the regenerative process T. If T can be simulated, σ could be estimated in a straightforward fashion. As we discussed in the last section, however, it is usually impossible to determine the transition function Q_k, so we must find an alternative approach to the problem. In section 4.4 we returned to the original process and partitioned the sample path into 'quasi-cycles' using an independent sequence of Bernoulli random variables. The question we address in this section is how to estimate the variance constant σ using this sequence of random variables. Before we being recall that $Z = (X, C)$ is an aperiodic $(\Omega', \varphi, \lambda, k)$-recurrent GSMOP with truncated distributions and transition function P. The process T is a regeenrative process based on a decomposition of P^k. Finally the process Q is the original process Z augmented by an independent sequence of Bernoulli random variables χ. Also recall that f is a function defined on Ω' by

$$f(x, c) = \min_{c_i > 0} c_i \cdot g(x)$$

for some measurable function $g: S \to R$.

The most naive approach to estimating σ using Q is to form the same statistics based on its quasi-cycles as one would ordinarily form for a regenerative process. Therefore, let

$$s_{11}^{\mathcal{Q}} = \frac{1}{n-1}\sum_{j=1}^{n}(L_j - \bar{L})^2,$$

$$s_{22}^{\mathcal{Q}} = \frac{1}{n-1}\sum_{j=1}^{n}(\upsilon_j - \bar{\upsilon})^2,$$

$$s_{12}^{\mathcal{Q}} = \frac{1}{n-1}\sum_{j=1}^{n}(L_j\upsilon_j - \bar{L}\bar{\upsilon}),$$

$$\hat{r}_{\mathcal{Q}} = \bar{L}/\bar{\upsilon},$$

$$s_{\mathcal{Q}}^2 = s_{11}^{\mathcal{Q}} - 2\hat{r}_{\mathcal{Q}}\, s_{12}^{\mathcal{Q}} + \hat{r}_{\mathcal{Q}}^2 s_{22}^{\mathcal{Q}}$$

where the $\{L_j\}$ and $\{\upsilon_j\}$ sequences are defined in 4.4.6. In all of these statistics n is a suppressed argument. To determine the asymptotic properties of these statistics, let us couple the process T with the sequence χ. Let $\psi = (T,\chi) = \{\psi_n = (T_n, \chi_n) : n = 0, 1, \ldots, \}$ be a process defined on $\Xi \times (0,1)$ with initial measure $\nu(A,\gamma) = \mathrm{P}(T_0 \in A)1_{\{1\}}(\gamma)$ and probability transition function

$$\mathrm{P}_\psi\{(v,\delta,\ell,y,\gamma),(A,\chi)\} = \mathrm{P}'\{(v,\delta,\ell,y),A\}1_{\{\gamma\}}(\chi)$$
$$\text{for } (v,\delta,\ell,y) \in \Xi, \ell = 2,3,\ldots,k$$
$$\mathrm{P}_\psi\{(v,\delta,1,y,\gamma),(A,\chi)\} = [p\chi + (1-p)(1-\chi)]\mathrm{P}'\{(v,\delta,1,y),A\},$$

where P′ is the transition function for T defined in 4.4.1.

Notice that ψ is the process T supplemented by an independent sequence of independent Bernoulli random variables. These random variables are generated at the same times and with the same parameter as the δ process in T. Recall that V (the first coordinate of T) behaves marginally according to the same probabilitic rules as the original process Z. By studying this coupled process we can make deductions about the asymptotic properties of the variance estimators of both T and $\mathcal{Q}$. The first step in comparing the asymptotic properties of s_T and $s_{\mathcal{Q}}$ is the comparison of the point estimates for $Eg\mathfrak{B}'$, $\hat{r}_T$ and $\hat{r}_{\mathcal{Q}}$.

(1) THEOREM. If $E|g(\mathfrak{B}')| < +\infty$, then $\lim_{n\to+\infty} \bar{r}_{\mathcal{Q}} = Eg\mathfrak{B}'$ with probability 1.

Proof. We know that $\bar{r}_T \to Eg\mathfrak{B}'$ with probability 1, and

$$P\left(\sum_{i=0}^{n} \min_{C_{ij}>0} C_{ij}\, g(X_i) \le x\right) = P\left(\sum_{i=0}^{n} \min_{K_{ij}>0} K_{ij}\, g(U_i) \le x\right)$$

for any measurable function f. Since convergence to a constant in probability is the same as convergence in distribution, a simple argument on the remainder terms and then a continuous mapping argument implies that the asymptotic properties of the estimators are the same. The details of the argument will be omitted. ∎

The remaining terms is $s^2_{\mathcal{Q}}$ are sample variance and covariance terms. We will examine *one of these terms*, $s^{\mathcal{Q}}_{11}$, in some detail. The other two can be handled by similar reasoning. To study the asymptotic behavior of $s^{\mathcal{Q}}_{11}$ we will now define a number of random variables related to the coupled process ψ. Let

$$\beta_0(\psi) = \beta_0(T) = \beta_0(\mathcal{Q}) = 0$$

$$\beta_i(\psi) = \min\{\, nk > \beta_{i-1}(\psi) \mid \chi_{nk} = 1, \delta_{nk} = 1 \,\}$$

$$\beta_i(T) = \min\{\, nk > \beta_{i-1}(T) \mid \delta_{nk} = 1 \,\}$$

$$\beta_i(\mathcal{Q}) = \min\{\, nk > \beta_{i-1}(\mathcal{Q}) \mid \chi_{nk} = 1 \,\}$$

$$O_n(\psi) = \max\{\, i \mid \beta_i(\psi) \le n \,\}$$

$$O_n(T) = \max\{\, i \mid \beta_i(T) \le n \,\}$$

$$O_n(\mathcal{Q}) = \max\{\, i \mid \beta_i(\mathcal{Q}) \le n \,\}.$$

The β sequences mark the beginning of the cycles in their respective sequences, while the O processes are the counting processes for the three sequences.

(2) **THEOREM.** If $\int \min_{c_j>0} c_j |g(x)|\, d\pi < \infty$, then

$$\lim_{n\to\infty} s_{11}^{\mathcal{Q}}$$

$$= \frac{1}{\mathrm{E}\{\, O_{\beta_1(\psi)}(\mathcal{Q})\,\}} E\left\{ \sum_{j=1}^{O_{\beta_1(\psi)}(\mathcal{Q})} (L_j - \mathrm{E}_\pi\{\, \beta_1(\mathcal{Q})\,\}\mathrm{E}_\pi\{\, f(Z_0)\,\})^2 \right\} \quad \mathrm{P}_\nu - a.e.$$

for any probability measure ν.

Note: (a) In this theorem we introduce the notion of a supercycle. A supercycle will be a piece of the coupled process ψ between consecutive $\beta_i(\psi)$s. Observe that these supercycles are legitimate regenerative cycles, since whenever a supercycle begins a cycle for the T process begins as well.

(b) This theorem says that the asymptotic variance estimate equals the expected sum of variances of the quasi-cycles in a supercycle divided by the expected number of quasi-cycles in a super cycle.

(c) To take advantage of the Strong Law of Large Numbers, we must find a way to break the $\mathcal{Q}$ sequence into independent and identically distributed block; this is one reason it is convenient to couple the χ process with the T process.

(d) The notation that arises in this theorem looks more intimidating than it really is. First $O_{\beta_n(\mathcal{Q})}(\psi)$ is the number of supercycles completed by the end of the n^{th} quasi-cycle. Then $\beta_{O_{\beta_n(\mathcal{Q})}(\psi)}(\psi)$ denotes the end of the last supercycle complete by the end of the n^{th} quasi-cycle. Finally $\Upsilon_n = O_{\beta_{O_{\beta_n(\mathcal{Q})}(\psi)}(\psi)}(\mathcal{Q})$ is the number of quasi-cycles which occur in those complete supercycles. Also, suppose that $\Upsilon'_n = O_{\beta_{O_{\beta_n(\mathcal{Q})}(\psi)}(\psi)+1}(\mathcal{Q})$.

Proof. Now $s_{11}^{\mathcal{Q}}$ has a limit if and only if

$$\lim_{n\to+\infty} \frac{1}{n-1} \sum_{i=1}^{n} (L_j - \mathrm{E}_\pi\{\, \beta_i(\mathcal{Q})\,\}\mathrm{E}_\pi\{\, f(Z_0)\,\})^2 \qquad (3)$$

exists. Furthermore, when the limits exist they will be the same. To see this note

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} f(Z_i) = \lim \frac{1}{n}\Big\{ \sum_{k=1}^{O_n(\mathfrak{a})} \sum_{i=\beta_{k-1}(\mathfrak{a})}^{\beta_k(\mathfrak{a})-1} f(Z_i) + \sum_{j=\beta_{O_n(\mathfrak{a})}(\mathfrak{a})}^{n} f(Z_i) \Big\} \tag{4}$$
$$= E_\pi\{ f(Z_0) \} \quad P_\nu - a.e.$$

for any probability measure ν, by the ratio limit theorem (Lemma 2.1.7). Now

$$\frac{1}{n}\sum_{i=\beta_{O_n(\mathfrak{a})}(\mathfrak{a})}^{n} f(Z_i) \le \frac{1}{n}\sum_{\beta_{O_n(\mathfrak{a})}(\mathfrak{a})}^{\beta_{O_n(\mathfrak{a})+1}(\mathfrak{a})} f(Z_i)$$
$$\le \frac{1}{n}\max_{s\in S} f(s)\,\upsilon_{O_n(\mathfrak{a})+1}$$

where υ_k is the length of the k^{th} quasi-cycle. The last expression above goes to zero as n goes to infinity since $\upsilon_{O_n(\mathfrak{a})+1} < +\infty$ with probability 1. Then from (4) we have

$$\lim_{n\to\infty} \frac{1}{O_n(\mathfrak{a})} \sum_{k=1}^{O_n(\mathfrak{a})} \sum_{i=\beta_{k-1}(\mathfrak{a})}^{\beta_k(\mathfrak{a})-1} f(Z_i) = \lim_{n\to\infty} \frac{n}{O_n(\mathfrak{a})} \lim_{n\to\infty} \frac{1}{O_n(\mathfrak{a})} \sum_{k=1}^{O_n(\mathfrak{a})} \sum_{i=\beta_{k-1}(\mathfrak{a})}^{\beta_k(\mathfrak{a})-1} f(Z_i)$$
$$= E_\pi\{ f(Z_0) \} \lim_{n\to\infty} \frac{n}{O_n(\mathfrak{a})}$$
$$= E_\pi\{ f(Z_0) \} E\{ \beta_{O_n(\mathfrak{a})+1} - \beta_{O_n(\mathfrak{a})} \}$$
$$= E\{ \beta_1(\mathfrak{a}) \} E_\pi\{ f(Z_0) \}$$

The next to last equality is true by the discrete version of the Elementary Renewal Theorem and the last is valid since the number of quasi-cycles in a supercycle is an i.i.d. sequence. Now we must find the limit of (3). This limit can be determined by breaking a sample path into blocks using the regenerative points of the coupled process. To do this notice

$$\frac{1}{n-1}\sum_{j=1}^{n}(L_j - \mathrm{E}_\pi\{\beta_1(\mathrm{Q})\}\mathrm{E}_\pi\{f(Z_0)\})^2$$

$$= \frac{1}{n}\sum_{j=1}^{\Upsilon_n}(L_j - \mathrm{E}_\pi\{\beta_1(\mathrm{Q})\}\mathrm{E}_\pi\{f(Z_0)\})^2$$

$$+ \sum_{j=\Upsilon_n+1}^{n}(L_j - \mathrm{E}_\pi\{\beta_1(\mathrm{Q})\}\mathrm{E}_\pi\{f(Z_0)\})^2.$$

We can show by the usual arguments that the second term goes to zero as n goes to infinity and we only need to evaluate the first term. To apply the Strong Law of Large Numbers let us multiply and divide by $O_{\beta_n(\mathrm{Q})}(\psi)$. There are then two limits to evaluate. From the Strong Law of Large Numbers and the fact that $O_{\beta_n(\mathrm{Q})}(\psi) \to +\infty$ as $n \to +\infty$, we have

$$\lim_{n\to\infty}\frac{1}{O_{\beta_n(\mathrm{Q})}(\psi)}\sum_{k=1}^{\Upsilon_n}(L_j - \mathrm{E}_\pi\{\beta_1(\mathrm{Q})\}\mathrm{E}_\pi\{f(Z_0)\})^2$$

$$= \mathrm{E}\sum_{j=\Upsilon_n}^{\Upsilon'_n}(L_j - \mathrm{E}_\pi\{\beta_i(\mathrm{Q})\}\mathrm{E}_\pi\{f(Z_0)\})^2 \quad \mathrm{P}_\nu - a.e.$$

for any probability measure ν.

The number of quasi-cycles in a supercycle is an i.i.d. sequence, so the limits of summation can be replaced by 1 and $O_{\beta_1(\psi)}(\mathrm{Q})$. We must also evaluate $\lim_{n\to\infty} O_{\beta_n(\mathrm{Q})}(\psi)/n$. This term represents the average number of supercycles in n quasi-cycles. This requires the same counting process reasoning that we used in Theorem 4.2.1. Bound $n/O_{\beta_n(\mathrm{Q})}(\psi)$ by $\beta_{O_{\beta_n(\mathrm{Q})}(\psi)}(\mathrm{Q})/O_{\beta_n(\mathrm{Q})}(\psi)$ and $\beta_{O_{\beta_n(\mathrm{Q})}(\psi)+1}(\mathrm{Q})/O_{\beta_n(\mathrm{Q})}(\psi)$ and use the Strong Law and the fact that $\beta_n(\mathrm{Q}) \to \infty$ as $n \to \infty$ to infer that

$$\lim_{n\to\infty}\frac{n}{O_{\beta_n(\mathrm{Q})}} = \mathrm{E}\{O_{\beta_k(\psi)}(\mathrm{Q}) - O_{\beta_{k-1}(\psi)}(\mathrm{Q})\} \quad \mathrm{P}_\nu - a.e.$$

for any probability measure ν. This expectation is the same for each super-cycle, so we may choose k=1. Therefore (2) is proven. ∎

Now we have demonstrated that s_{11}^{Q} has a limit. It is possible to write an analogous expression for s_{11}^{T} merely by replacing Q with T and L with Y everywhere. We would like to show that the limits are these same. To pursue this goal requires a variation of Wald's Lemma.

(5) THEOREM. If $X_1, X_2, \ldots$ are identically distributed random variables having finite expectations, and if N is a positive integer-valued random variable independent of the X_n's with $EN < +\infty$, then

$$\mathrm{E}\sum_{i=1}^{N} X_i = \mathrm{E}\{N\}\mathrm{E}\{X_1\}.$$

Proof. Let

$$Y_n = \begin{cases} 1, & \text{if } N \geq n, \\ 0, & \text{if } N < n. \end{cases}$$

Then we have

$$\sum_{n=1}^{N} X_n = \sum_{n=1}^{\infty} X_n Y_n$$

and

$$\mathrm{E}\sum_{n=1}^{N} X_n = \mathrm{E}\sum_{n=1}^{\infty} X_n Y_n = \sum_{n=1}^{\infty} \mathrm{E}(X_n Y_n). \tag{6}$$

If $X_n Y_n$ is replaced by $|X_n Y_n|$, the interchange in (6) is justified because all the terms are positive. This implies that the original interchange is allowable by Lebesgue's Dominated Convergence Theorem. The random variable Y_n is

independent of X_n, so we obtain

$$\mathrm{E}\sum_{n=1}^{N} X_n = \sum_{n=1}^{\infty} \mathrm{E}\{X_n\}\mathrm{E}\{Y_n\}$$

$$= \mathrm{E}\{X_n\}\sum_{n=1}^{\infty} \mathrm{E}\{Y_n\}$$

$$= \mathrm{E}\{X_1\}\sum_{n=1}^{\infty} \mathrm{P}(N \geq n)$$

$$= \mathrm{E}\{X_1\}\mathrm{E}\{N\}. \quad \blacksquare$$

We are now ready to show that the limits of s_{11}^{T} and $s_{11}^{\mathfrak{Q}}$ are the same. Recall from Lemma 2.1.5 that we can choose k so that the GSMOP Z is $(\Omega', \pi, \lambda, k)$-recurrent.

(7) THEOREM. Let Z be an aperiodic $(\Omega', \varphi, \lambda, k)$-recurrent GSMOP which has invariant measure π. Suppose that Z is also $(\Omega', \pi, \lambda', k)$-recurrent. If $\int |f(z)|\, d\pi < \infty$, then

$$\frac{1}{\mathrm{E}\{O_{\beta_1(\psi)}(\mathfrak{Q})\}}\mathrm{E}\left\{\sum_{j=1}^{O_{\beta_1(\psi)}(\mathfrak{Q})} (L_j - \mathrm{E}\{\beta_1(\mathfrak{Q})\}\mathrm{E}_\pi\{f(Z_0)\})^2\right\} =$$

$$\frac{1}{\mathrm{E}\{O_{\beta_1(\psi)}(T)\}}\mathrm{E}\left\{\sum_{j=1}^{O_{\beta_1(\psi)}(T)} (Y_j - \mathrm{E}\{\beta_1(T)\}\mathrm{E}_\pi\{f(V_0)\})^2.\right\}$$

Proof. First notice that δ and χ are both sequences of Bernoulli random variables with parameter p and the roles they play in determining the $\beta(\psi)$ sequence are symmetric; therefore,

$$\mathrm{E}\{O_{\beta_1(\psi)}(\mathfrak{Q})\} = \mathrm{E}\{O_{\beta_1(\psi)}(T)\}.$$

We only need to show the equivalence of the sums. Suppose initially that $\varphi = \pi$, the limiting distribution of Z. In this case the terms in each sum are identically distributed, the number of terms in each sum is independent of each summand, and all the relevant expectations are finite, so applying Theorem (5) in each expectation, we have

$$\mathrm{E}\left\{ \sum_{j=1}^{O_{\beta_1(\psi)}(\mathcal{Q})} (L_j - \mathrm{E}\{\beta_1(\mathcal{Q})\}\mathrm{E}_\pi\{f(Z_0)\})^2 \right\}$$

$$= \mathrm{E}\{O_{\beta_1(\psi)}(\mathcal{Q})\}\,\mathrm{E}\{(L_j - \mathrm{E}\{\beta_1(\mathcal{Q})\}\,\mathrm{E}_\pi\{f(Z_0)\})^2\}$$

and

$$\mathrm{E}\left\{ \sum_{j=1}^{O_{\beta_1(\psi)}(T)} (Y_j - \mathrm{E}\{\beta_1(T)\}\mathrm{E}_\pi\{f(V_0)\})^2 \right\}$$

$$= \mathrm{E}\{O_{\beta_1(\psi)}(T)\}\mathrm{E}\{(Y_j - \mathrm{E}\{(\beta_1(T)\}\mathrm{E}_\pi\{f(V_0)\})^2\}.$$

As we noted above $\mathrm{E}\{O_{\beta_1(\psi)}(\mathcal{Q})\} = \mathrm{E}\{O_{\beta_1(\psi)}(T)\}$, so we must only show the equivalence of the second terms. We know, however, that $\mathrm{E}\{\beta_1(T)\} = \mathrm{E}\{\beta_1(\mathcal{Q})\}$ and $\mathrm{E}_\pi\{f(Z_0)\} = \mathrm{E}\{f(V_0)\}$, so we must only show that $\mathrm{E}\{Y_j\} = \mathrm{E}\{L_j\}$ and $\mathrm{E}\{Y_j^2\} = \mathrm{E}\{L_j^2\}$. We can apply Theorem (4) again to each side of the first equation to see that

$$\mathrm{E}\{Y_j\} = \mathrm{E}\sum_{n=0}^{\beta_1-1} f(V_n) = \mathrm{E}\{\beta_1(T) - 1\}\mathrm{E}_\pi\{(V_0)]$$

and

$$\mathrm{E}\{L_j\} = \mathrm{E}\sum_{n=0}^{\beta_1(\mathcal{Q})-1} f(Z_n) = \mathrm{E}\{\beta_1(\mathcal{Q}) - 1\}\mathrm{E}_\pi\{f(Z_0)\}.$$

The random variables $\beta_1(\mathcal{Q})$ and $\beta_1(T)$ are both determined by Bernoulli sequences with the same parameter and the second terms are equal since both V

and Z are Markov chains with the same initial distribution, transition function and limiting random variable.

Now we will examine the second moments of the cycles of T and the quasi-cycles. Suppose

$$D_j(\mathfrak{Q}) = \begin{cases} 1, & \text{if } \beta_1(\mathfrak{Q}) \geq j, \\ 0, & \text{if } \beta_1(\mathfrak{Q}) < j. \end{cases}$$

Then we can write

$$\Big(\sum_{i=0}^{\beta_1(\mathfrak{Q})-1} \min_{C_{ij}>0} C_{ij} f(X_i)\Big)^2 = \Big(\sum_{i=0}^{\infty} f(Z_i) D_i(\mathfrak{Q})\Big)^2.$$

Thus

$$\mathrm{E}\left\{ \sum_{j=0}^{\beta_1(\mathfrak{Q})-1} \min_{C_{ij}>0} C_{ij} f(X_j)^2 \right\} = \mathrm{E}\left\{ \Big(\sum_{i=0}^{\infty} f(Z_i) D_i(\mathfrak{Q})\Big)^2 \right\}$$

$$= \sum_{i=0}^{\infty} \mathrm{E}(f(Z_i) D_i(\mathfrak{Q}))^2$$

$$+ 2 \sum_{0 \leq i_1 < i_2 < \infty} E\{ f(Z_{i_1}) D_{i_1}(\mathfrak{Q}) f(Z_{i_2}) D_{i_2}(\mathfrak{Q}) \}.$$

The interchange is justified if $f(Z_i)$ is replace by its absolute value throughout, since all the terms would then be positive. This implies that the original interchange is allowable by Lebesgue's Dominated Convergence Theorem. Now $D_j(\mathfrak{Q})^2 = D_j(\mathfrak{Q})$ and is independent of X_i for each i. Thus we have

$$\mathrm{E}\left\{\left(\sum_{i=0}^{\beta_1(\mathbb{Q})-1} f(Z_i)\right)^2\right\}$$

$$= \sum_{i=0}^{\infty} \mathrm{E}(f(Z_i)D_i(\mathbb{Q}))^2$$

$$+2 \sum_{0\le i_1<i_2<\infty} E\{ f(Z_{i_1})D_{i_1}(\mathbb{Q})f(Z_{i_2})D_{i_2}(\mathbb{Q})\}$$

$$= \sum_{i=0}^{\beta_1(\mathbb{Q})} \mathrm{E}(f(Z_i))^2 \mathrm{E}(D_i(\mathbb{Q}))$$

$$+2 \sum_{0\le i_1<i_2<\infty} \mathrm{E}\{ D_{i_1}(\mathbb{Q})D_{i_2}(\mathbb{Q})\}\mathrm{E}\{ f(Z_{i_1})f(Z_{i_2})\}$$

$$= \mathrm{E}_\pi\{ (f(Z_0))^2\}\,\mathrm{E}\{ \beta_1(\mathbb{Q})-1\}$$

$$+2 \sum_{0\le i_1<i_2<\infty} \mathrm{E}\{ D_{i_1}(\mathbb{Q})D_{i_2}(\mathbb{Q})\}\mathrm{E}\{ f(Z_{i_1})f(Z_{i_2})\}.$$

Similarly $\mathrm{E}(Y_j^2)$ can be decomposed. But we note that

$$\mathrm{E}_\pi\{ (f(Z_0))^2\} = \mathrm{E}_\pi\{ (f(V_0))^2\},$$
$$\mathrm{E}\{ \beta_1(T)\} = \mathrm{E}\{ \beta_1(\mathbb{Q})\},$$
$$\mathrm{E}\{ D_{i_1}(\mathbb{Q})D_{i_2}(\mathbb{Q})\} = \mathrm{E}\{ D_{i_1}(T)D_{i_2}(T)\},$$
$$\mathrm{E}\{ f(Z_{i_1})f(Z_{i_2})\} = \mathrm{E}\{ f(V_{i_1})f(V_{i_2})\}.$$

The final step in the proof is to establish that we do not need the initial distribution of the chain to be π in order to claim that the theorem is true. Recall that Theorem 5 demonstrated that the limits for $s_{11}^{\mathbb{Q}}$ was true P_ν—a.e. for any initial measure ν. The terms in the limit do not depend on the initial measure. The theorem proof is complete. ∎

Section 4.6. AN EXAMPLE

THE THEORY DEVELOPED in Sections 4.3 and 4.4 proposed a technique for finding point estimates and confidence intervals for functions of the equilibrium distribution of a GSMO whose associated GSMOP Z is Doeblin recurrent. It is, however, likely to be troublesome for a practitioner to determine the modified densities required to develop a truncated chain for Z if it is Harris rather than Doeblin recurrent. The purpose of this section is to test the techniques proposed on a model which is Harris recurrent.

The model we will examine is the (2,2)-server cyclic queue with feedback and four customers. A schematic diagram of the systems is presented in Figure 1. As we noted in Section 3.1, there is not a single set, since at least two servers are active at all times. The system is aperiodic and (A, φ, λ, k)-recurrent for some choices of A, φ, λ and k.

The first step is to determine the recurrence parameters which are satisfied by the system. In Section 4.2 we demonstrated that chains with densities positive on $[0, \infty)$ were recurrent with A chosen as the subset of Ω' which has all its clocks less than some U. If we choose U sufficiently large, the sample path we generate using the original process should be virtually indistinguishable from the sample path we would generate if we truncated the densities at U. To determine what U is appropriate we must know the distributions associated with our model.

In order to facilitate the calculation of theoretical means and variances, we suppose that each server has a gamma distribution. Let both servers at station A have a $\Gamma(2,1)$ distribution; at station B let one server have a $\Gamma(2,1)$ distribution, the other a $\Gamma(3,1)$ distribution. For these distributions a relatively small choice of U is sufficient. If U=20, for example, $P(\Gamma(2,1) > 10) = .9999949$ and $P(\Gamma(3,1) > 10) = .9999968$.

Now we must choose λ and k. Unfortunately, how we choose these parameters is largely a matter of guesswork. It is, of, course, possible to be very conservative in our choices, letting $k=100$ and $\lambda = .001$, for example. Such choices will do little for the efficiency of our simulation, however.

We can be guided by a few facts, however. Our treatment of the asymptotic properties of our variance estimate required that the Markov chain be π-recurrent for our choice of k. Also the ergodic condition gives us some help for sets with large φ-measure. If $\sup_{x,y,A}|P(x,A) - P(y,A)| \leq 1-\epsilon$, then Doeblin's condition is satisfied with parameters $(P^k(x,\cdot), k, 1-\epsilon/2, \epsilon/2)$, which means that when $P^k(x,A) > 1-\epsilon/2$ then $P^k(y,A) > \epsilon/2$ for all choices of y.

Choosing k such that Z is π-recurrent for k or determining appropriated ϵ which satisfies the coefficient of ergodicity for a given k is not straightforward, however. In order to reflect our ignorance about which k and ϵ to choose, we have simulated the system with a number of different choices for each. We have chosen $k=10$, 20 and 30 with $\lambda = .4$, .6, .9 and .99. We would like to choose λ as large as possible for a given k since this will increase the number of regenerations in a run of fixed length.

The functions we choose to test the method are indicators which can be used to determine the equilibrium distribution of system. As we pointed out in Section 2.2 the state space of this system can be described by a pair to indicate which servers at station A are busy (x_1, x_2), a pair to indicate which servers at station B are busy (x_3, x_4), and a coordinate (x_5) to indicate how many jobs are in service at station A. Let us suppose that there are four jobs in the system. The functions we will test are $f(x) = 1_{\{i\}}(x_5)$ for $i=1,2,3,4$. Tables 1 and 2 contain the 12 combinations of λ and k for f_1. Tables 3 and 4 contain the results for f_2, 5 and 6 contain f_3 and 7 and 8 contain f_4.

Let the feedback loop be chosen with probability .5 at each completion of a service by one of the servers at station A. Under these assumptions the

network can be viewed as a Markov chain with 78 states and the theoretical means and variances can be calculated.

In all the tables that follow, P.E. means the point estimate for f, and H.L. denotes the half length of the confidence interval. Again we will compute 90 percent confidence intervals.

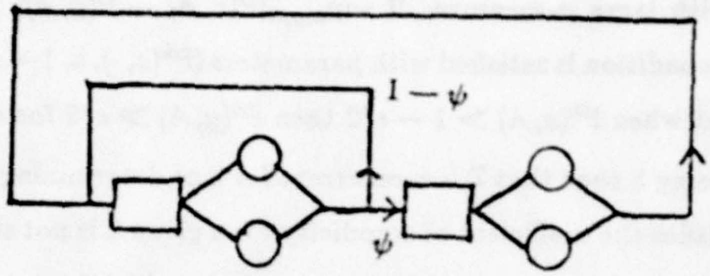

Figure 1. Two Server Cyclic Queue with Feedback

TABLE 1: SIMULATION OF TWO SERVER CYCLIC QUEUE WITH FEEDBACK

4 jobs – servers at A - $\Gamma(2,1), \Gamma(2,1)$ – servers at B - $\Gamma(2,1), \Gamma(3,1)$ – 90 percent confidence

$f_1(x) = 1_{\{1\}}(x_5)$									
$\lambda = .4$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{a}/\bar{a}$	P.E.	H.L.	$\hat{a}/\bar{a}$	P.E.	H.L.	$\hat{a}/\bar{a}$
100	.0949	.0165	.1847	.1092	.0128	.2031	.1046	.0096	.1993
200	.0948	.0120	.1769	.1056	.0088	.2014	.1032	.0085	.1799
300	.0996	.0010	.2015	.1039	.0074	.2115	.1014	.0053	.1737
400	.1052	.0091	.2226	.1027	.0062	.1992	.1037	.0048	.1897
500	.1074	.0.0082	.2222	.1054	.0057	.2253	.1041	.0037	.1901
THEORY	.1044		.1986	.1044		.1986	.1044		.1986

$f_1(x) = 1_{\{1\}}(x_5)$									
$\lambda = .6$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{a}/\bar{a}$	P.E.	H.L.	$\hat{a}/\bar{a}$	P.E.	H.L.	$\hat{a}/\bar{a}$
100	.1008	.0207	.1957	.1021	.0113	.1250	.1153	.0128	.2145
200	.1101	.0151	.1969	.1034	.0097	.1780	.1080	.0082	.1925
300	.1020	.0116	.1813	.1000	.0080	.1767	.1037	.0067	.1899
400	.1029	.0103	.1881	.1028	.0072	.1913	.1066	.0061	.2030
500	.1007	.0091	.1812	.1052	.0068	.2099	.1045	.0055	.2014
THEORY	.1044		.1988	.1044		.1986	.1044		.1986

TABLE 2: SIMULATION OF TWO SERVER CYCLIC QUEUE WITH FEEDBACK

4 jobs – servers at A - $\Gamma(2,1), \Gamma(2,1)$ – servers at B - $\Gamma(2,1), \Gamma(3,1)$ – 90 percent confidence

$f_1(x) = 1_{\{1\}}(x_5)$									
$\lambda = .9$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$
100	.1115	.0239	.1714	.0909	.0183	.1912	.1152	.0158	.2263
200	.1107	.0173	.1782	.0911	.0121	.1694	.1153	.0108	.2095
300	.1078	.0147	.1951	.0915	.0106	.1896	.1065	.0081	.1764
400	.1048	.0126	.1863	.0951	.0090	.1870	.1065	.0071	.1792
500	.1025	.0111	.1793	.0956	.0070	.1837	.1037	.0037	.1901
THEORY	.1044		.1986	.1044		.1986	.1044		.1986

$f_1(x) = 1_{\{1\}}(x_5)$									
$\lambda = .99$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$
100	.1115	.0274	.1980	.1098	.0197	.2106	.0893	.0140	.1559
200	.1092	.0179	.1700	.1032	.0128	.1756	.0987	.0101	.1614
300	.1140	.0148	.1749	.1024	.0105	.1771	.1034	.0090	.1929
400	.1119	.0130	.1794	.1073	.0094	.1869	.0986	.0076	.1854
500	.1119	.0116	.1792	.1059	.0085	.1914	.0997	.0069	.1914
THEORY	.1044		.1983	.1044		.1986	.1044		.1986

TABLE 3: SIMULATION OF TWO SERVER CYCLIC QUEUE WITH FEEDBACK

4 jobs – servers at A - $\Gamma(2,1), \Gamma(2,1)$ – servers at B - $\Gamma(2,1), \Gamma(3,1)$ – 90 percent confidence

$f_2(x) = 1_{\{2\}}(x_5)$									
$\lambda = .4$	$k = 10$			$k = 20$			$k = 30$		
CYCLES	P.E.	H.L.	$\hat{o}/\bar{a}$	P.E.	H.L.	$\hat{o}/\bar{a}$	P.E.	H.L.	$\hat{o}/\bar{a}$
100	.2503	.0189	.2560	.2393	.0164	.3367	.2477	.0177	.2518
200	.2497	.0127	.2361	.2414	.0107	.2958	.2446	.0074	.2364
300	.2376	.0127	.2361	.2377	.0091	.3252	.2427	.0063	.2467
400	.2358	.0108	.2590	.2388	.0075	.2940	.2446	.0055	.2526
500	.2361	.0084	.2403	.2414	.0066	.2838	.2440	.0049	.2538
THEORY	.2420		.2757	.2420		.2757	.2420		.2757

$f_2(x) = 1_{\{2\}}(x_5)$									
$\lambda = .6$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{o}/\bar{a}$	P.E.	H.L.	$\hat{o}/\bar{a}$	P.E.	H.L.	$\hat{o}/\bar{a}$
100	.2577	.0264	.2938	.2391	.0175	.3014	.2616	.0128	.2073
200	.2585	.0185	.2996	.2376	.0119	.2691	.2474	.0099	.2821
300	.2569	.0149	.2863	.2414	.0099	.2751	.2456	.0079	.2610
400	.2402	.0135	.3188	.2419	.0087	.2762	.2452	.0069	.2618
500	.2401	.0116	.2917	.2430	.0078	.2753	.2428	.0062	.2642
THEORY	.2420		.2757	.2420		.2757	.2420		.2757

TABLE 4: SIMULATION OF TWO SERVER CYCLIC QUEUE WITH FEEDBACK

4 jobs – servers at A - $\Gamma(2,1), \Gamma(2,1)$ – servers at B - $\Gamma(2,1), \Gamma(3,1)$ – 90 percent confidence

$f_2(x) = 1_{\{2\}}(x_5)$									
$\lambda = .9$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$
100	.2480	.0279	.2312	.2490	.0199	.2248	.2312	.0164	.2438
200	.2467	.0214	.2703	.2490	.0199	.2248	.2312	.0124	.2438
300	.2380	.0165	.2450	.2476	.0144	.2462	.2470	.0099	.2657
400	.2365	.0143	.2401	.2506	.0106	.2697	.2429	.0085	.2567
500	.2350	.0130	.2493	.2495	.0096	.2759	.2427	.0078	.2722
THEORY	.2420		.2757	.2420		.2757	.2420		.2722

$f_2(x) = 1_{\{2\}}(x_5)$									
$\lambda = .99$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$
100	.2319	.0327	.2841	.2406	.0226	.2709	.2270	.0180	.2606
200	.2090	.0216	.2505	.2442	.0166	.2915	.2307	.0126	.2525
300	.2204	.0173	.2409	.2382	.0131	.2731	.2365	.0106	.2694
400	.2242	.0154	.2541	.2330	.0112	.2726	.2336	.0093	.2770
500	.2310	.0137	.2516	.2372	.0099	.2649	.2392	.0084	.2826
THEORY	.2420		.2757	.2420		.2757	.2420		.2757

TABLE 5: SIMULATION OF TWO SERVER CYCLIC QUEUE WITH FEEDBACK

4 jobs – servers at A - $\Gamma(2,1),\Gamma(2,1)$ – servers at B - $\Gamma(2,1),\Gamma(3,1)$ – 90 percent confidence

$f_3(x) = 1_{\{3\}}(x_5)$									
$\lambda = .4$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$
100	.3615	.0232	.3301	.3564	.0144	.2582	.3756	.0118	.2877
200	.3649	.0159	.3353	.3574	.0105	.2838	.3679	.0084	.2888
300	.3648	.0122	.3215	.3606	.0081	.2566	.3720	.0076	.3603
400	.3648	.0105	.3082	.3612	.0069	.2567	.3731	.0065	.3479
500	.3694	.0096	.3189	.3601	.0065	.2827	.3728	.0059	.3561
THEORY	.3671		.3120	.3671		.3120	.3671		.3120

$f_3(x) = 1_{\{3\}}(x_5)$									
$\lambda = .6$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$
100	.3563	.0238	.2517	.3648	.0166	.2465	.3796	.0145	.2797
200	.3614	.0163	.2468	.3643	.0118	.2578	.3677	.0102	.2822
300	.3679	.0139	.2605	.3653	.0100	.2845	.3704	.0086	.2930
400	.3737	.0128	.2893	.3684	.0089	.2881	.3683	.0076	.2997
500	.3715	.0115	.2868	.3693	.0079	.2818	.3675	.0071	.3179
THEORY	.3671		.3120	.3671		.3120	.3671		.3120

TABLE 6: SIMULATION OF TWO SERVER CYCLIC QUEUE WITH FEEDBACK

4 jobs – servers at A - $\Gamma(2,1), \Gamma(2,1)$ – servers at B - $\Gamma(2,1), \Gamma(3,1)$ – 90 percent confidence

$f_3(x) = 1_{\{3\}}(x_5)$									
$\lambda = .9$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{\sigma}/\overline{a}$	P.E.	H.L.	$\hat{\sigma}/\overline{a}$	P.E.	H.L.	$\hat{\sigma}/\overline{a}$
100	.3508	.0313	.3183	.3673	.0216	.2970	.3714	.0185	.2976
200	.3566	.0225	.3176	.3602	.0151	.2926	.3676	.0138	.3278
300	.3449	.0181	.3059	.3684	.0122	.2871	.3648	.0109	.3121
400	.3488	.0151	.2828	.3668	.0107	.2905	.3639	.0096	.3210
500	.3507	.0134	.2779	.3678	.0097	.2932	.3692	.0089	.3395
THEORY	.3671		.3120	.3671		.3120	.3671		.3120

$f_3(x) = 1_{\{3\}}(x_5)$									
$\lambda = .99$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{\sigma}/\overline{a}$	P.E.	H.L.	$\hat{\sigma}/\overline{a}$	P.E.	H.L.	$\hat{\sigma}/\overline{a}$
100	.3591	.0330	.3009	.3693	.0265	.3833	.3507	.0172	.2402
200	.3500	.0243	.3229	.3719	.0181	.3539	.3533	.0124	.2527
300	.3521	.0194	.3053	.3712	.0145	.3362	.3641	.0111	.3045
400	.3533	.0167	.2993	.3727	.0124	.3305	.3698	.0095	.2957
500	.3613	.0150	.3050	.3694	.0112	.3397	.3682	.0081	.3061
THEORY	.3671		.3120	.3671		.3120	.3671		.3120

TABLE 7: SIMULATION OF TWO SERVER CYCLIC QUEUE WITH FEEDBACK

4 jobs – servers at A - $\Gamma(2,1), \Gamma(2,1)$ – servers at B - $\Gamma(2,1), \Gamma(3,1)$ – 90 percent confidence

	$f_4(x) = 1_{\{4\}}(x_5)$								
$\lambda = .4$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$
100	.2716	.0298	.5461	.2792	.0200	.4994	.2707	.0164	.5492
200	.2845	.0207	.5689	.2679	.0144	.5401	.2634	.0110	.4912
300	.2903	.0162	.5688	.2642	.0120	.5614	.2637	.0091	.5054
400	.2792	.0143	.5725	.2767	.0105	.5951	.2671	.0077	.4945
500	.2739	.0128	.5655	.2698	.0091	.5617	.2688	.0070	.5012
THEORY	.2664		.5503	.2664		.5503	.2664		.5503

	$f_4(x) = 1_{\{4\}}(x_5)$								
$\lambda = .6$	$k = 10$			$k = 20$			$k = 30$		
Cycles	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$	P.E.	H.L.	$\hat{\sigma}/\bar{a}$
100	.2885	.0366	.5949	.2747	.0258	.5939	.2540	.0168	.3753
200	.2959	.0237	.5211	.2756	.0173	.5575	.2615	.0140	.5308
300	.2989	.0194	.5076	.2717	.0137	.5339	.2605	.0111	.4915
400	.2821	.0169	.5069	.2675	.0119	.5358	.2623	.0099	.5010
500	.2832	.0158	.5457	.2713	.0109	.5381	.2672	.0091	.5281
THEORY	.2664		.5503	.2664		.5503	.2664		.5503

TABLE 8: SIMULATION OF TWO SERVER CYCLIC QUEUE WITH FEEDBACK

4 jobs – servers at A - $\Gamma(2,1),\Gamma(2,1)$ – servers at B - $\Gamma(2,1),\Gamma(3,1)$ – 90 percent confidence

$f_4(x) = 1_{\{4\}}(x_5)$									
$\lambda = .9$		$k = 10$			$k = 20$			$k = 30$	
Cycles	P.E.	H.L.	$\hat{\sigma}/\overline{a}$	P.E.	H.L.	$\hat{\sigma}/\overline{a}$	P.E.	H.L.	$\hat{\sigma}/\overline{a}$
100	.2652	.0431	.6004	.2731	.0372	.5203	.2516	.0251	.5493
200	.2812	.0300	.5651	.2801	.0283	.5309	.2664	.0177	.5463
300	.2848	.0253	.6021	.2711	.0229	.5418	.2799	.0147	.5702
400	.2833	.0216	.5767	.2688	.0187	.5405	.2788	.0127	.5558
500	.2814	.0193	.5679	.2689	.0159	.5417	.2766	.0114	.5630
THEORY	.2664		.5503	.2664		.5503	.2664		.5503

$f_4(x) = 1_{\{4\}}(x_5)$									
$\lambda = .99$		$k = 10$			$k = 20$			$k = 30$	
Cycles	P.E.	H.L.	$\hat{\sigma}/\overline{a}$	P.E.	H.L.	$\hat{\sigma}/\overline{a}$	P.E.	H.L.	$\hat{\sigma}/\overline{a}$
100	.2710	.0434	.5196	.2842	.0293	.4692	.2488	.0263	.5579
200	.2946	.0327	.5857	.2474	.0199	.4266	.2819	.0184	.5581
300	.2873	.0259	.5437	.2563	.0165	.4375	.2771	.0149	.5516
400	.2856	.0222	.5341	.2600	.0146	.4589	.2745	.0131	.5519
500	.2890	.0197	.5223	.2620	.0132	.4714	.2699	.0097	.5531
THEORY	.2664		.5503	.2664		.5503	.2664		.5503

Before analyzing the results of our simulations, let us consider the results we would expect to obtain and the effects of varying λ and k. First, for our functions, point estimates are obtained by dividing the portion of the simulation run the function is positive by the length of the run. This is the same technique that is used in practically every method for estimating $E\{f(\mathfrak{X}')\}$. Therefore we would expect our point estimates to be close to the exact value, regardless of the choice of k and λ.

The values of k and λ should effect the variance calculation, however. In specifying λ and k, we are hypothesizing that

$$P^k(x, A) \geq \lambda\varphi(A) \tag{1}$$

for all x and measurable A. If we choose $\varphi = \pi$, for example, the geometric convergence of iterates of P to π (see Lemma 2.1.17) leads us to believe that larger values of k admit more values of λ which satisfy (1). This implies that the larger the value of k, the more values of λ will yield acceptable estimates of σ. On the other hand, smaller choices of λ should also yield more accurate estimates of σ since smaller values of k would satisfy (1) for a given λ. Both effects (choosing smaller λ and larger k) tend to lengthen each regenerative cycle and thus lengthen the 'real' time a simulation of a fixed number of cycles requires.

Now let us turn to the results of the simulation. As we expected, the coverage of the true value of the estimated parameter was good, and was fairly uniform over various choices of λ and k. The true value of $E\{f\mathfrak{X}\}$ was covered in 11 of 12 runs for $\lambda = .4, .9$ and .99 and in 10 of 12 runs for $\lambda = .6$. Overall, the true value was covered in 92 percent of the runs. The coverage was also fairly uniform over k. For $k = 10$, there was coverage in 13 of 16 runs; for $k = 20$, 14 of 16 runs, and for $k = 30$, 16 of 16 runs.

Now let us consider the variance estimates the simulations obtained. As

we expected, the larger the value of k the better the estimate. The improvement is particularly dramatic for $k = 30$. Of the 16 runs with $k = 30$, 13 estimated the variance constant with an error of less than 5 percent. For $k = 20$, 7 of 16 runs had an error of less than 5 percent, 6 mor had an error between 5 and 10 percent. For $k = 10$, 5 and 7 were the corresponding numbers. This data seems to confirm our intuition about the effect of k.

On the other hand, varying the choice of λ produced puzzling results. For $\lambda = .6, .9$ and $.99$, the results tend to confirm our intuitions. For the runs for $\lambda = .4$, unexpected results occurred. Of the 9 runs (of 48 total) in which the variance estimated differed from the true value of σ by more than 10 percent, 5 occurred with $\lambda = .4$. We are considering a small sample, but nevertheless, this is somewhat surprising.

Finally, it is interesting to note that the results were good for $k = 10$ for large values of λ. We would not expect that $P^{10}(A) \geq .99\pi(A)$, yet the results for this choice of parameters was very good.

AD-A075 549 STANFORD UNIV CA DEPT OF OPERATIONS RESEARCH F/G 12/1
SIMULATING GENERALIZED SEMI-MARKOV PROCESSES.(U)
SEP 79 L D FOSSETT N00014-76-C-0578
UNCLASSIFIED TR-50 NL

2 OF 2
AD-A075549

END
DATE FILMED
11-79
DDC

CHAPTER V

CONCLUSIONS

As the use of stochastic systems becomes more extensive in the study of complex phenomena, the need for a theoretical understanding of these models and a practical means of analyzing them becomes more acute. We have proposed an approach which demonstrates the inherent regenerative structure of many of these stochastic systems and capitalizes on that structure in a way that allows us to estimate many quantities conerned with the behavior of the system in the 'long run'. Our task now to reexamine the approach and evaluate its strengths and weaknesses.

The major weaknesses of the approach are obvious. Certainly the assumptions we have made about the system under study have been stricter than we would like. The most important of these are the requirement that S be finite and the assumption that a GSMO without a single set must be Doeblin recurrent in order to apply the estimation procedure in Sections 4.4 and 4.5. S may be infinite (in an open network, for example) and the densities of the clocks may have support on the entire positive half line in many situations of interest. In practice, these difficulties may be circumvented by restricting S and truncating the densities in such a way that the system spends only a tiny portion of time outside this set. Also the example in Section 4.6 seems to indicate that it is not necessary to explicitly truncate the distributions.

Nevertheless, it is theoretically preferable not to require these approximation procedures.

The chief practical difficulty lies in choosing λ and k. At present the choice is largely guesswork. It is interesting, however, that many choices seem to give good results —even when we would guess that the parameters do not satisfy the recurrence condition.

The strengths of our approach are equally clear. By demonstrating that GSMOPs are frequently recurrent, we can use the Athreya-Ney-Nummelin construction to find regenerative processes that are closely related to the original GSMOP. This allows us to say, at least theoretically, a great deal about the process of interest. Furthermore, with only a small amount of additional effort, we can find point estimates and confidence intervals for functions of the process when it is equilibrium, even when we cannot explicitly determine a sample path for a companion regenerative process. The procedures we propose seem to be applicable to a large variety of stochastic systems, including many queueing networks.

A study of this kind can never answer all questions that arise from the development of a technique and we will now consider a few that remain unanswered by our discussion. First we would like to determine which of our assumptions can be relaxed. The three extensions we would like to make are: (1) allowing S to be countably infinite—this generalization might require a good deal of extra work, we frequently used the fact that S was finite; (2) allowing speeds other than 1 in GSMOs without single sets—this would probably lead to a notational nightmare; and (3) requiring a GSMO to be only Harris recurrent in order to use our independent sequences procedure—this generalization might be obtainable.

Second, in order to make the method accessible to the practitioner, we must improve our means of choosing λ and k. Although the procedure seems

fairly robust, it is clearly unsatisfactory to have only one's intuition about the network to guide the selection of these parameters.

Finally, this technique is designed to provide ways of estimating functions of the equilibrium distribution of a GSMP. In the simulation of queueing networks, for example, it is often desireable to make inferences about the response time (the time required for a job to traverse some portion of the network). This is usually considered to be a separate question but it would be interesting to determine if the companion regenerative processes to a network could be used to extend to a more general setting the existing results in this area.

BIBLIOGRAPHY

1. Athreya, K.B. and P. Ney. (1977). A new approach to the limit theory of Markov chains. Preprint.

2. Athreya, K.B. and P. Ney. (1978). A new approach to the limit theory of Markov chains. *Trans. Amer. Math. Soc.* 245, 493–501.

3. Billingsley, P. (1968). *Convergence of Probability Measures*, John Wiley, New York.

4. Breiman, L. (1968). *Probability*. Addison–Wesley, Reading, Mass.

5. Chung, K.L. (1967). *Markov Chains with Stationary Transition Probabilities*. Springer–Verlag, Berlin.

6. Chung, K.L. (1974). *A Course in Probability Theory*. Academic Press, New York.

7. Çinlar, E. (1975). *Introduction to Stochastic Processes*. Prentice–Hall, Englewood Cliffs, New Jersey.

8. Crane, M.A. and D.L. Iglehart. (1974a). Simulating stable stochastic systems, I: General multiserver queues. *J. Assoc. Comput. Mach.*, 21, 103–113.

9. Crane, M.A. and D.L. Iglehart. (1974b). Simulating stable stochastic systems, II: Markov chains," *J. Assoc. Comput. Mach.* 21, 114–123.

10. Crane, M.A. and D.L. Iglehart. (1975). Simulating stable stochastic systems, III: Regenerative processes and discrete event simulations. *Operat. Res.* 23, 33–45.

11. Crane, M.A. and A.J. Lemoine. (1978) *An Introduction to the Regenerative Method for Simulation*. Springer–Verlag, Berlin.

12. Dobrušin, R.L. (1956). Central limit theorems for nonstationary Markov

chains. *Theor. Probability Appl.* 1, 65–80, 329–383.

13. Doeblin, W. (1940). Éléments d'une théorie générale des chaînes simples constantes de Markoff," *Ann. Sci. École Norm. Sup.* bf 57, 61–111.

14. Doob, J.L. (1953). *Stochastic Processes.* John Wiley, New York.

15. Feller, W. (1971). *An Introduction to Probability Theory.* II John Wiley, New York.

16. Gaver, D.P. and G.S. Shedler. (1971). Control variable methods in the simulation of a multiprogrammed computer system. *Naval Res. Logist. Quart.* 18, 435–450.

17. Harris, T.E. (1956). The existence of stationary measures for certain Markov processes. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability.* II, Jerzy Neymann, ed., University of California Press, Berkeley.

18. Hordijk, A. and R. Schassberger. (1976) Weak convergence for generalized semi-Markov processes. Unpublished manuscript.

19. Iosifescu M. and R. Theodorescu. (1969). *Random Processes and Learning.* Springer–Verlag, Berlin.

20. Lavenberg S.S. and G.S. Shedler. (1976). Stochastic modeling of processor scheduling with applications to data base management systems. *IBM J. Res. Develop.* 20, 437–448.

21. Matthes, K. (1962) Zur Theorie der Bedienungsprozesse. *Trans. 3rd Prague Conference on Inf. Thy. Stat. Dec. Fct.* Prague.

22. Neveu, J. (1964) *Bases Mathématiques du Calcul des Probabilitiés,* Masson et C^{ie}, Paris.

23. Nummelin, E. (1978). A splitting technique for Harris recurrent Markov chains. *Z. Wahrsheinlichkeitstheorie verw. Gebiete.* 43, 309–318.

24. Orey, S. (1971). *Limit Theorems for Markov Chain Transition Probabilities,* Van Nostrand, New York.

25. Pearson, K. (1957). *Tables of the Incomplete Γ Function.* Cambridge University Press, London.

26. Revuz, D. (1975). *Markov Chains*, North-Holland, Amsterdam.

27. Rudin, W.R. (1974). *Real and Complex Analysis*, McGraw-Hill, New York.

28. Royden, H.L. (1968) *Real Analysis*. MacMillan, New York.

29. Schassberger, R. (1975). Insensitivity of steady-state distributions of generalized semi-Markov processes: I, II. Department of Mathematics, University of Calgary.

30. Smith, W.L. (1955). Regenerative stochastic processes. *Proc. Roy. Soc. London Ser. A.* 232, 6–31.

30. Whitt, W. (1976). Continuity of generalized semi-Markov processes. School of Organization and Management, Yale University.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 50	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): SIMULATING GENERALIZED SEMI-MARKOV PROCESSES	5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL REPORT
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): LAWRENCE D. FOSSETT	8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0578
9. PERFORMING ORGANIZATION NAME AND ADDRESS: DEPARTMENT OF OPERATIONS RESEARCH, STANFORD UNIVERSITY, STANFORD, CA 94305	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 042-343
11. CONTROLLING OFFICE NAME AND ADDRESS: STATISTICS AND PROBABILITY PROGRAM, OFFICE OF NAVAL RESEARCH (Code 436), ARLINGTON, VA 20360	12. REPORT DATE: SEPTEMBER 1979
	13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

GENERALIZED SEMI-MARKOV PROCESSES, NETWORKS OF QUEUES, RECURRENT MARKOV PROCESSES, REGENERATIVE METHOD, SIMULATION, STATISTICAL ESTIMATION

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE REVERSE SIDE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

One approach to modeling queueing networks and other complex stochastic systems which has received some attention in the literature is the generalized semi-Markov process (GSMP). This idea is an example of the supplementary variables approach to non-Markovian systems. This approach 'supplements' the natural description of the system by variables which contain information about the past history of the system. In this way, a model of a non-Markovian system can be made Markovian. For GSMPs the supplementary variables are clocks which record the amount of time until the occurrence of various events which could influence the system. In a queueing network, for example, each server and each arrival stream would be associated with a clock. By including these clock readings as part of the description of the system, only the present state of the system is required to predict future behaviour. This means that the new model is Markovian and therefore amenable to analysis via the use of Markov chain theory. To use these processes for simulation purposes a central limit theorem is required. Obtaining this result is the second goal of this paper. Our approach to this problem is to find closely related regenerative processes on which to base the central limit theorem for the process under study. New results in the theory of Markov chains on a general state space make it clear how these regenerative processes can be constructed.